Um estudo sobre otimização de funções reais de várias variáveis: teoria e aplicações

Priscila Cordero Leal

SERVIÇO DE PÓS-GRADUAÇÃO DO ICMC-USP

Data de Depósito:

Assinatura: ______________________

Priscila Cordero Leal

Um estudo sobre otimização de funções reais de várias variáveis: teoria e aplicações

Dissertação apresentada ao Instituto de Ciências Matemáticas e de Computação – ICMC-USP, como parte dos requisitos para obtenção do título de Mestra em Ciências – Programa de Mestrado Profissional em Matemática. *VERSÃO REVISADA*

Área de Concentração: Matemática

Orientadora: Profa. Dra. Katia Andreia Gonçalves de Azevedo

USP – São Carlos
Abril de 2017

Ficha catalográfica elaborada pela Biblioteca Prof. Achille Bassi
e Seção Técnica de Informática, ICMC/USP,
com os dados fornecidos pelo(a) autor(a)

Leal, Priscila Cordero
L634u Um estudo sobre otimização de funções reais de
várias variáveis: teoria e aplicações / Priscila
Cordero Leal; orientadora Katia Andreia Gonçalves
de Azevedo. -- São Carlos -- SP, 2017.
101 p.

Dissertação (Mestrado - Programa de Pós-graduação
em Mestrado Profissional em Matemática em Rede
Nacional) -- Instituto de Ciências Matemáticas e
de Computação, Universidade de São Paulo, 2017.

1. Otimização. 2. Máximos e Mínimos. 3. Cálculo de
várias variáveis. I. Azevedo, Katia Andreia Gonçalves
de, orient. II. Título.

Priscila Cordero Leal

A study on optimization of real functions of several variables: theory and applications

Master dissertation submitted to the Instituto de Ciências Matemáticas e de Computação – ICMC-USP, in partial fulfillment of the requirements for the degree of Mathematics Professional Master's Program. *FINAL VERSION*

Concentration Area: Mathematics

Advisor: Profa. Dra. Katia Andreia Gonçalves de Azevedo

USP – São Carlos
April 2017

Dedico este trabalho às pessoas que sempre estão ao meu lado, transmitindo-me fé, amor, alegria, determinação, paciência e coragem, tornando meus dias mais felizes e bonitos;

À minha mãe maravilhosa Neide Cordero Crisol;

Aos irmãos mais que especiais Aletéia Cordero Leal, Murilo Cordero Leal e Milena Cordero Leal;

À minha linda sobrinha Helena Leal Oliveira;

À minha amada avó Maria de Lourdes Crisol Donha;

Ao meu avô Francisco Cordero Donha, que mesmo distante se faz sempre presente.

AGRADECIMENTOS

Agradeço a Deus pelo dom da vida e por todas as oportunidades de me tornar cada vez melhor e estar sempre em busca do bem;

À minha maravilhosa orientadora Katia Andreia Gonçalves de Azevedo pela disposição em me ajudar e por sempre me incentivar, me fazendo acreditar sempre mais em minha capacidade;

À minha mãe Neide Cordero Crisol e à minha irmã Milena Cordero Leal, pessoas excepcionais, que estiveram presentes em todos os momentos dessa jornada, me ajudando e me apoiando para que meu objetivo fosse alcançado com sucesso;

Aos meus amigos Janayna Resende Barbosa, Talles Eduardo Nazar Cerizza e Rafael de Freitas Manço pelo companherismo em todos os momentos vivenciados até aqu;

A todos que fazem parte de minha história, contribuindo direta ou indiretamente, para meu desenvolvimento.

À CAPES.

"Aqueles que contam com o Senhor renovam suas forças;
Ele dá-lhes asas de águia.
Correm sem se cansar, vão para a frente sem se fatigar."
(Isaías 40:31)

RESUMO

PRISCILA CORDERO LEAL. **Um estudo sobre otimização de funções reais de várias variáveis: teoria e aplicações**. 2017. 101 p. Dissertação (Mestrado em Ciências – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Otimizar significa determinar estratégias para se obter as melhores alternativas em busca dos objetivos traçados. Em matemática, otimização refere-se ao estudo de problemas em que se deseja maximizar ou minimizar uma determinada função através da escolha sistemática dos valores de variáveis dentro de um conjunto viável. Um dos métodos de se determinar os máximos e mínimos de funções é utilizando-se o cálculo em várias variáveis, o qual será abordado nesse trabalho. Possíveis situações relacionadas ao cotidiano são apresentadas para que o processo de otimização possa ser abordado por alunos do Ensino Médio.

Palavras-chave: Otimização, Máximos e Mínimos, Cálculo de várias variáveis.

ABSTRACT

PRISCILA CORDERO LEAL. **A study on optimization of real functions of several variables: theory and applications**. 2017. 101 p. Dissertação (Mestrado em Ciências – Programa de Mestrado Profissional em Matemática) – Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo, São Carlos – SP, 2017.

Optimizing means determing strategies to obtain better alternatives in the search of the stated aims. In Mathematics, optimization refers to the study of problems in which we want to maximize or to minimize a certain function through a systematic choice of variable values in a viable set. One of the methods used to determine the maximums and the minimums of functions is using calculus in several variables, what will be discussed in this work. Possible situations related to daily life are presented so that the optimization process can be studied by high school students.

Keywords: Optimization, Maximums and Minimums, Calculus in several variables.

LISTA DE ILUSTRAÇÕES

SUMÁRIO

INTRODUÇÃO

Em tempos atuais, em busca do desenvolvimento, de aprimoramento, o termo otimizar aparece com enorme destaque.

A Matemática é uma ferramenta poderosa no estudo de processos de otimização e através dela vem-se criando métodos que possibilitam a tomada de melhores decisões.

Aproximar situações reais, que são consideradas ou discutidas no dia-a-dia, dos alunos em sala de aula, tem-se mostrado uma forma bastante interessante em situações educacionais. Sabendo-se das dificuldades com relação à Matemática apresentadas pelos alunos ao longo dos anos escolares, este tema pode ser bastante explorado em aula como fator motivador.

Para esses processos de otimização, é de suma importância criar uma função a ser trabalhada, a qual denominamos função objetivo, porém em diversas situações nos deparamos com situações limitadoras, dessa forma criam-se novas funções, denominadas funções restrições.

O método que utilizamos no decorrer desse trabalho é baseado no cálculo diferencial com várias variáveis. Dessa maneira, fizemos um estudo com conceitos preliminares utilizados no cálculo, para depois estudarmos os métodos de otimização, ou seja, métodos pelos quais conseguimos encontrar valores que maximizam ou minimizam os modelos matemáticos a serem estudados.

No primeiro capítulo, desenvolvemos os pré-requisitos necessários para compreender os conceitos, as hipóteses e as demonstrações que foram apresentadas ao longo do trabalho. Estudos sobre diferenciabilidade, regra da cadeia e polinômio de Taylor para funções de várias variáveis a valores reais foram abordados.

No segundo capítulo, abordamos alguns tópicos sobre a teoria de otimização para funções de várias variáveis a valores reais divididos em problemas de otimização condicionados e não condicionados. Alguns problemas que desenvolvemos são de fácil interpretação e provenientes de situações reais.

Para alcançar os alunos do Ensino Básico, com esta teoria, faz-se necessário o estudo de ferramentas e conteúdos que são normalmente abordados em cursos universitários, como por exemplo, derivação. Entretanto, citamos exemplos em que podemos trabalhar o conceito de máximo e mínimo de uma função, levando em consideração outros conhecimentos já adquiridos, ao menos, por alunos do Ensino Médio. Estes exemplos são abordados no terceiro capítulo.

Todas as figuras apresentadas foram feitas por meio do software Mathematica, porém podem ser utilizados outros softwares, como por exemplo, o Geogebra.

CAPÍTULO

1

CONCEITOS PRELIMINARES

Neste capítulo revisaremos alguns conceitos de cálculo em váriais variáveis que serão importantes no desenvolvimento dos próximos capítulos. Em geral, assumiremos que o leitor tenha uma certa familiaridade com o conceito de função e com o espaço euclideano n-dimensional, ver [1]

Através de funções de várias variáveis pode-se modelar uma infindável quantidade de fenômenos dos mais diversos ramos, por exemplo, o volume V de um cilindro é dado por $V = \pi r^2 h$, onde r é o raio e h é a altura; a função de Cobb-Douglas, que modela o desempenho de uma economia, dada por $q = kx^{\alpha_1} y^{\alpha_2}$, onde q é o produto, x é o capital, y é a mão-de-obra e α_1 e α_2 são constantes apropriadas, entre vários outros.

1.1 Função de Várias Variáveis

Definição 1. Seja $A \subset \mathbb{R}^n$. Uma função f de A em $\mathbb{R}$ é uma regra que associa a cada $x \in A$ um único número real $f(x)$. O conjunto A será chamado de domínio de f e o conjunto $\{f(x) : x \in A\}$ será chamado de imagem de f. A variável x é chamada de variável independente e a variável $f(x)$ é chamada de variável dependente.

O gráfico de uma função é definido como o conjunto $Gr(f) = \{(x, f(x)) \in \mathbb{R}^{n+1} / x \in A\}$.

Observação 1. No que segue, usaremos a seguinte notação para denotar uma função como descrita na Definição (1)

$$f : A \subset \mathbb{R}^n \to \mathbb{R}$$

$$x \mapsto f(x).$$

Se $n = 2$, denotamos as variáveis independentes por (x, y) e a função por $z = f(x, y)$.

Se $n = 3$, denotamos as variáveis independentes por (x, y, z) e a função por $w = f(x, y, z)$, e assim, sucessivamente.

Exemplo 1. Considerando a função $f(x, y) = \sqrt{y - x} + \sqrt{1 - y}$, temos que seu domínio é dado pelo conjunto de todos os pares $(x, y) \in \mathbb{R}^2$, com $y - x \geq 0$ e $1 - y \geq 0$, ou seja $D_f = \{(x, y) \in \mathbb{R}^2 | y \geq x \text{ e } y \leq 1\}$ e sua imagem é o conjunto dos números reais positivos. As Figuras 1 e 2 mostram, respectivamente, o domínio e o gráfico de f.

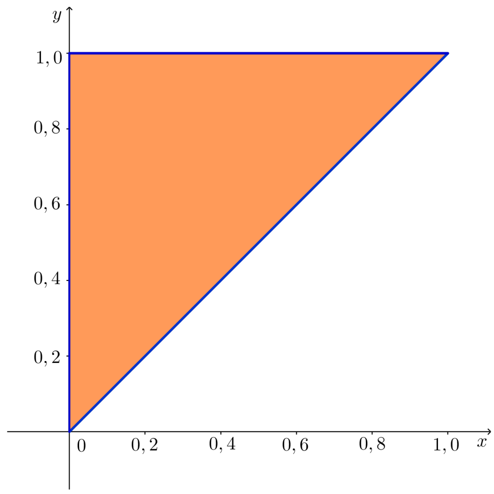

Figura 1 – Domínio de $f(x, y) = \sqrt{y - x} + \sqrt{1 - y}$.

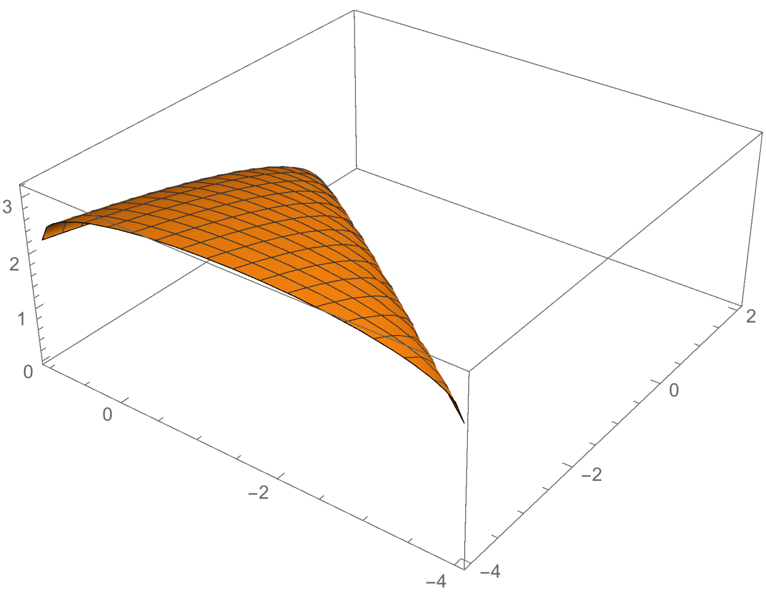

Figura 2 – Gráfico de $f(x, y) = \sqrt{y - x} + \sqrt{1 - y}$.

Exemplo 2. A função $f(x, y) = x^2 + y^2$ tem como domínio todo o $\mathbb{R}^2$ e imagem o conjunto dos números reais não negativos.A Figura 3 mostra o gráfico de f.

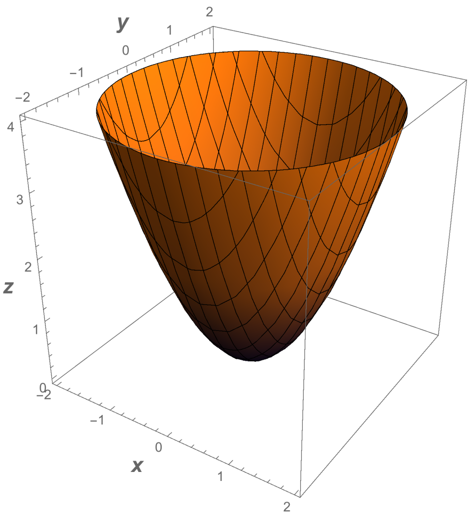

Figura 3 – Gráfico de $f(x,y) = x^2 + y^2$.

Definição 2. Sejam $z = f(x,y)$ uma função e c pertencente ao conjunto imagem de f. O conjunto de todos os pontos (x,y) pertencentes ao domínio de f tais que $f(x,y) = c$ denomina-se curva de nível de f correspondente ao nível $z = c$. Assim f é constante sobre cada curva de nível. O gráfico de f é um subconjunto do $\mathbb{R}^3$. Uma curva de nível é um subconjunto do domínio de f, portanto, do $\mathbb{R}^2$.

Exemplo 3. Para a função do Exemplo 2, as curvas de nível, no nível $c \in Imf$, são circunferências com centro na origem e raio $\sqrt{c}$, como podemos verificar na Figura 4.

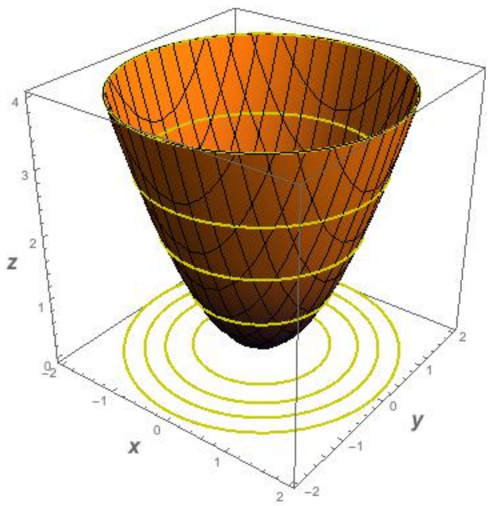

Figura 4 – Gráfico da função $f(x,y) = x^2 + y^2$ e a representação de suas curvas de nível no plano xy.

1.2 Funções diferenciáveis

O principal objetivo desta seção é o estudo de alguns conceitos básicos de funções diferenciáveis ou deriváveis. Para começar, introduziremos algumas notações e lembraremos do conceito de função contínua e da derivada de uma função de uma varíavel real a valores reais.

1.2.1 Diferenciabilidade para funções reais a valores reais

Definição 3. Sejam $f : A \subset \mathbb{R} \to \mathbb{R}$ uma função e $x_0 \in A$ um ponto de acumulação de A. Dizemos que f é contínua em x_0 quando $\lim\limits_{x \to x_0} f(x) = f(x_0)$. Dizemos que f é contínua se f for contínua em cada ponto de seu domínio.

Exemplo 4. Como exemplo de funções contínuas podemos citar as funções polinomiais, as funções racionais, as funções modulares, etc.

Exemplo 5. A função $f(x) = \begin{cases} x \operatorname{sen}\left(\dfrac{1}{x}\right), \; x \neq 0 \\ 0, \; x = 0 \end{cases}$, é contínua em todo o seu domínio.

Definição 4. Sejam $f : A \subset \mathbb{R} \to \mathbb{R}$ uma função e $x_0 \in A$ um ponto de acumulação. Definimos a derivada de f em x_0, denotada por $f'(x_0)$, por

$$f'(x_0) = \frac{df}{dx}(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

desde que o limite exista.

Podemos interpretar geometricamente o significado de f ser derivável em um ponto. Seja $C = Gr(f)$. Para definirmos a reta tangente, basta conhecer um ponto $P(x_0, f(x_0))$ e seu coeficiente angular m, então a equação da reta tangente r à curva em P é dada por $y - f(x_0) = m(x - x_0)$. Mas como obter m para que r seja tangente à curva em P? Consideremos um outro ponto Q, arbitrário sobre a curva, cujas coordenadas são $(x_0 + h, f(x_0 + h))$. A reta que contém P e Q é denominada reta secante à curva.

Analisemos agora a variação do coeficiente angular da reta secante fazendo Q se aproximar de P, ou seja, tomando h cada vez menor, conforme Figuras 5, 6 e 7 onde na primeira delas, P e Q estão distantes entre si, já na segunda figura estão mais próximas e na terceira a distância entre eles é bem pequena.

Podemos observar que se P está próximo de Q, o coeficiente angular da reta secante, m_{PQ} deve estar próximo do coeficiente angular m da reta r, ou seja, o coeficiente angular da secante tem um limite m quando Q tende a P, que é o coeficiente angular da reta tangente r.

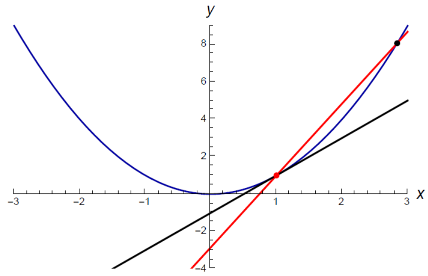

Figura 5 – Aproximação da reta tangente por retas secantes, com pontos P e Q distantes entre si.

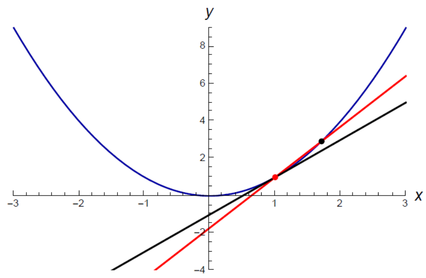

Figura 6 – Aproximação da reta tangente por retas secantes, com pontos P e Q mais próximos em relação à figura 5.

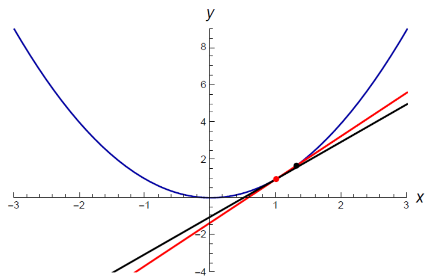

Figura 7 – Aproximação da reta tangente por retas secantes, com pontos P e Q bem próximos.

Indicando a abscissa de Q por $x = x_0 + h$, assim $h = x$ - x_0, e sabendo que a abscissa de P é x_0, então se $Q \to P$, temos que $h \to 0$, o que equivale a $x \to x_0$.

Assim,

$$m = \lim_{x \to x_0} m_{PQ} = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0},$$

(se este limite existe), é o coeficiente angular da reta tangente r e denotamos tal limite por $f'(x_0)$.

Logo, a derivada da função em um ponto fornece o coeficiente angular da reta tangente ao gráfico desta função, neste ponto.

Exemplo 6. Seja $f(x) = x$, provaremos que $f'(x_0) = 1$, para todo $x_0 \in \mathbb{R}$.

De fato, $f'(x_0) = \lim\limits_{h\to 0} \dfrac{f(x_0+h)-f(x_0)}{h} = \lim\limits_{h\to 0} \dfrac{x_0+h-x_0}{h} = \dfrac{h}{h} = 1.$

Podemos definir a função derivada $f' : D \subset \mathbb{R} \to \mathbb{R}$ dada por $f'(x) = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{h}$, onde D é um subconjunto de A tal que, se $x \in D$, tal limite existe.

No exemplo anterior, $f'(x) = 1$, $\forall x \in \mathbb{R}$.

Exemplo 7. Se $f, g : \mathbb{R} \to \mathbb{R}$ são dadas por $f(x) = sen(x)$ e $g(x) = cos(x)$, então, suas funções derivadas são:

$f'(x) = \lim\limits_{h\to 0} \dfrac{f(x+h)-f(x)}{h} = \lim\limits_{h\to 0} \dfrac{sen(x+h)-sen(x)}{h} =$

$\lim\limits_{h\to 0} \dfrac{2sen\left(\dfrac{x+h-x}{2}\right) cos\left(\dfrac{x+h+x}{2}\right)}{h} = \lim\limits_{h\to 0} \dfrac{2sen\left(\dfrac{h}{2}\right) cos\left(x+\dfrac{h}{2}\right)}{h}$, fazendo $\dfrac{h}{2} = t$, temos:

$f'(x) = \lim\limits_{t\to 0} \dfrac{2sen(t)cos(x+t)}{2t} = \lim\limits_{t\to 0} \dfrac{sen(t)}{t} \lim\limits_{t\to 0} cos(t+x) = 1\ cos(x) = cos(x)$, $\forall x \in \mathbb{R}$.

$g'(x) = \lim\limits_{h\to 0} \dfrac{g(x+h)-g(x)}{h} = \lim\limits_{h\to 0} \dfrac{cos(x+h)-cos(x)}{h}$

$= \lim\limits_{h\to 0} \dfrac{-2sen\left(\dfrac{x+h+x}{2}\right) sen\left(\dfrac{x+h-x}{2}\right)}{h} = \lim\limits_{h\to 0} \dfrac{-2sen\left(x+\dfrac{h}{2}\right) sen\left(\dfrac{h}{2}\right)}{h}$,

fazendo $\dfrac{h}{2} = t$, temos:

$g'(x) = \lim\limits_{t\to 0} \dfrac{-2sen(x+t)sen(t)}{2t} = \lim\limits_{t\to 0} -sen(t+x) \lim\limits_{t\to 0} \dfrac{sen(t)}{t} = -1\ sen(x) = -sen(x)$, $\forall x \in \mathbb{R}$.

O conceito de diferenciabilidade de uma função está associado à ideia de "suavidade"de seu gráfico.

Vamos entender este conceito inicialmente para funções de uma variável real.

Se $f : A \subset \mathbb{R} \to \mathbb{R}$ é uma função derivável, então seu gráfico é uma curva suave, sem pontos angulosos. Se f tem derivada em $x = x_0$, isto é, existe $f'(x_0)$, então:

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = f'(x_0)$$

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} - f'(x_0) = 0$$

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} = 0$$

$$\lim_{x \to x_0} \frac{f(x) - [f(x_0) + f'(x_0)(x - x_0)]}{x - x_0} = 0 \, .$$

Chamando $T(x) = f(x_0) + f'(x_0)(x - x_0)$ e $R(x) = f(x) - T(x)$, como $\lim\limits_{x \to x_0} \frac{R(x)}{x - x_0} = 0$ isto significa que $T(x)$ se aproxima de $f(x)$ (quando $x \to x_0$) “mais rápido”que x se aproxima de x_0. Em outras palavras, o gráfico de T, que é a reta tangente ao gráfico de f em $(x_0, f(x_0))$ é uma boa aproximação de f perto de x_0.

Observemos na Figura 8, a proximidade entre os gráficos de $f(x) = x^2$ e $T(x) = f(x_0) + f'(x_0)(x - x_0)$, com $x_0 = 1$, ou seja, $T(x) = 2x - 1$.

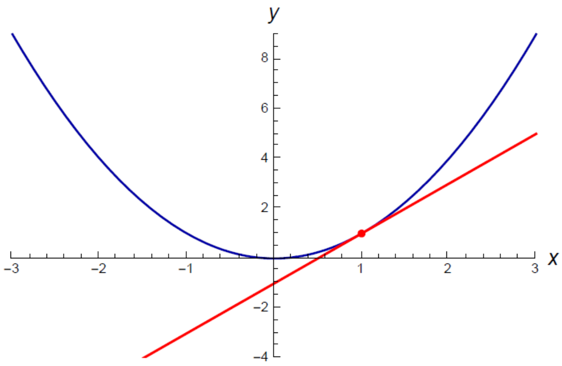

Figura 8 – Reta tangente ao gráfico de f em $(x_0, f(x_0))$.

Por outro lado, vemos que a reta tangente é a única reta com esta propriedade, isto é, se procurarmos por outra reta $y = m(x - x_0) + f(x_0)$ que quando $x = x_0$ vale $f(x_0)$ e satisfaz

$$0 = \lim_{x \to x_0} \frac{R(x)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - y}{x - x_0}$$

$$= \lim_{x \to x_0} \frac{f(x) - [m(x - x_0) + f(x_0)]}{x - x_0} = \lim_{x \to x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - m \right),$$

vemos, pela unicidade do limite, que $m = f'(x_0)$.

Dessa forma, se f tem derivada em x_0, tem reta tangente ao gráfico de f em $x = x_0$ e esta é uma boa aproximação para o gráfico de f, isto é, podemos aproximar o valor de f pelo valor de T, próximo a x_0:

$$\underbrace{f(x) - f(x_0)}_{\Delta y} = f'(x_0)(x - x_0) + R(x).$$

Observamos, também, que se $f : A \subset \mathbb{R} \to \mathbb{R}$ é diferenciável em x_0, então f é contínua em x_0.

De fato, se f é derivável em x_0, $\lim\limits_{x\to x_0} \dfrac{f(x)-f(x_0)}{x-x_0}$ existe e é igual a $f'(x_0)$. Precisamos provar que f é contínua em x_0, isto é, que $\lim\limits_{x\to x_0} f(x) = f(x_0)$.

Temos que,

$$f(x)-f(x_0) = \frac{f(x)-f(x_0)}{x-x_0}(x-x_0),\ x\neq x_0,$$

daí,

$$\lim_{x\to x_0}(f(x)-f(x_0)) = \lim_{x\to x_0}\frac{f(x)-f(x_0)}{x-x_0}\ .\ \lim_{x\to x_0}(x-x_0) = f'(x_0).0 = 0.$$

Ou seja,

$$\lim_{x\to x_0}(f(x)-f(x_0)) = 0$$

e, portanto,

$$\lim_{x\to x_0} f(x) = f(x_0).$$

1.2.2 Diferenciabilidade para funções de duas variáveis reais a valores reais

Para generalizar o conceito de diferenciabilidade para funções de duas variáveis gostaríamos que uma função de duas variáveis possuísse um "plano tangente"como uma boa aproximação e que também fosse contínua no ponto onde é diferenciável. Para isso, precisamos definir, inicialmente, o conceito de continuidade para funções de várias variáveis e o conceito de derivada parcial.

Definição 5. Suponha que $f : A\subset\mathbb{R}^2\to\mathbb{R}$ seja uma função e que $(x_0,y_0)\in A$ seja um ponto de acumulação de A. Dizemos que f é contínua em (x_0,y_0) , se e somente se, $\lim\limits_{(x,y)\to(x_0,y_0)} f(x,y) = f(x_0,y_0)$. Dizemos que f é contínua se f for contínua em cada ponto de seu domínio.

Exemplo 8. A função $f(x,y) = \begin{cases} \dfrac{x^2-y^2}{x^2+y^2}, & \text{se } (x,y)\neq(0,0) \\ 0, & \text{se } (x,y)=(0,0) \end{cases}$ é contínua para todo $(x,y)\neq(0,0)$, mas não é contínua em $(0,0)$, pois considerando os caminhos $\vec{r}_1(t) = (t,0)$ e $\vec{r}_2(t) = (0,t)$, $t\in\mathbb{R}$ temos,

$$\lim_{t\to 0} f(\vec{r}_1(t)) = \lim_{t\to 0} f(t,0) = \lim_{t\to 0}\frac{t^2}{t^2} = 1$$

$$\lim_{t\to 0} f(\vec{r}_2(t)) = \lim_{t\to 0} f(0,t) = \lim_{t\to 0} \frac{-t^2}{t^2} = -1.$$

Logo, $\nexists \lim\limits_{(x,y)\to(0,0)} f(x,y)$ e portanto, f não é contínua em $(0,0)$.

Vamos, agora, definir o conceito de derivada parcial. Considere $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, e $(x_0,y_0) \in A$. Fazendo $y = y_0$, podemos, então, considerar a função g de uma variável real dada por $g(x) = f(x,y_0)$. Olhando g como uma função de uma variável, podemos definir a derivada de g em x_0, $g'(x_0)$, por:

$$g'(x_0) = \lim_{x\to x_0} \frac{g(x)-g(x_0)}{x-x_0} = \lim_{x\to x_0} \frac{f(x,y_0)-f(x_0,y_0)}{x-x_0}.$$

Fazendo $x - x_0 = h$,

$$g'(x_0) = \lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h} = \frac{\partial f}{\partial x}(x_0,y_0),$$

chamada derivada parcial de f com respeito a x, no ponto (x_0,y_0), se o limite existir.

Geometricamente, a derivada parcial de f com respeito a x em (x_0,y_0), $\dfrac{\partial f}{\partial x}(x_0,y_0)$, é a inclinação da reta tangente no ponto $(x_0,y_0,f(x_0,y_0))$ à curva obtida pela intersecção do gráfico da função com um plano perpendicular ao plano xy, passando por $y = y_0$.

Exemplo 9. Seja a função $f(x,y) = x^2 - y^2$, cujo gráfico pode ser visto na Figura 9. A derivada parcial de f em relação a x é dada por $\dfrac{\partial f}{\partial x}(x_0,y_0) = 2x$ e a derivada parcial de f em relação a y dada por $\dfrac{\partial f}{\partial y}(x_0,y_0) = 2y$. A Figura 10 ilustra o plano perpendicular ao plano xy, passando por $y = y_0$, enquanto que a Figura 11 ilustra o plano perpendicular a plano xy, passando por $x = x_0$.

Considerando todos os pontos onde é possível calcular $\dfrac{\partial f}{\partial x}$, podemos definir a função

$$\frac{\partial f}{\partial x} : D_1 \subset A \subset \mathbb{R}^2 \to \mathbb{R}$$

$$(x,y) \mapsto \frac{\partial f}{\partial x}(x,y) = f_x(x,y)$$

chamada função derivada parcial de primeira ordem de f, com respeito à primeira variável.

Para calcular $\dfrac{\partial f}{\partial x}(x,y)$ basta observarmos que ela é a derivada da função $g(x) = f(x,y)$, mantendo y constante.

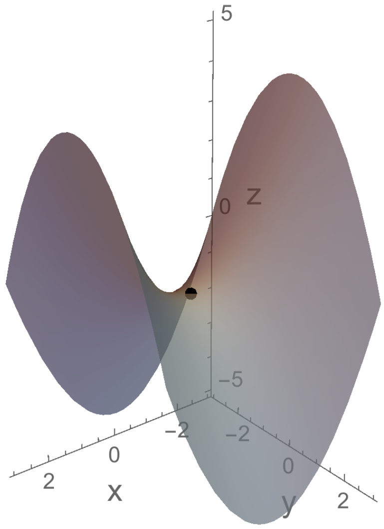

Figura 9 – Gráfico de $f(x,y) = x^2 - y^2$.

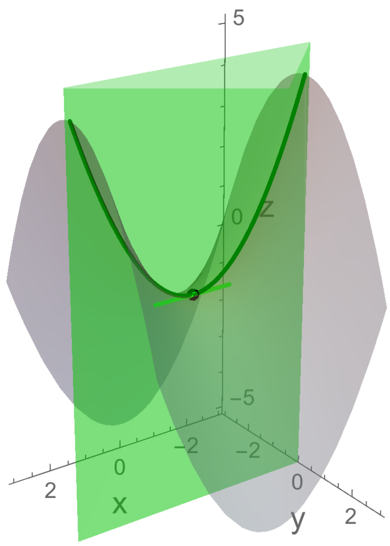

Figura 10 – Derivada parcial de f com respeito a x.

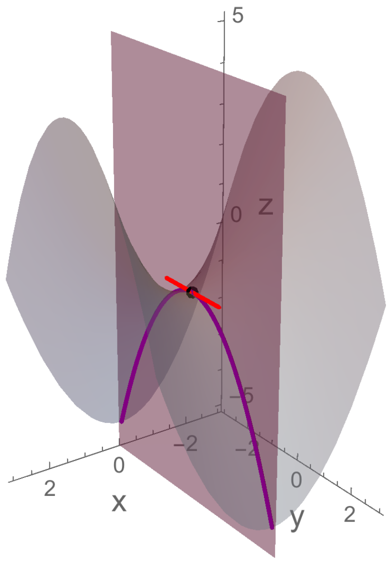

Figura 11 – Derivada parcial de f com respeito a y.

Exemplo 10. A derivada parcial em relação a variável x da função $f(x,y) = x^3 + x^2y^3 - 2y^2$ é obtida pela definição,

$$
\begin{aligned}
\frac{\partial f}{\partial x}(x_0,y_0) &= \lim_{x\to x_0} \frac{f(x,y_0) - f(x_0,y_0)}{x - x_0} \\
&= \lim_{x\to x_0} \frac{(x^3 + x^2y_0^3 - 2y_0^2) - (x_0^3 + x_0^2y_0^3 - 2y_0^2)}{x - x_0} \\
&= \lim_{x\to x_0} \frac{x^3 - x_0^3 + y_0^3(x^2 - x_0^2)}{x - x_0} \\
&= \lim_{x\to x_0} \frac{(x - x_0)(x^2 + xx_0 + x_0^2) + y_0^3(x - x_0)(x + x_0)}{x - x_0} \\
&= \lim_{x\to x_0} x^2 + xx_0 + x_0^2 + y_0^3(x + x_0) \\
&= x_0^2 + x_0^2 + x_0^2 + y_0^3(x_0 + x_0) \\
&= 3x_0^2 + 2x_0y_0^3.
\end{aligned}
$$

Analogamente, podemos definir $\dfrac{\partial f}{\partial y}(x_0,y_0)$. Fixando x_0 e definindo $h(y) = f(x_0,y)$, temos:

$$h'(y_0) = \lim_{y\to y_0} \frac{h(y) - h(y_0)}{y - y_0} = \lim_{y\to y_0} \frac{f(x_0,y) - f(x_0,y_0)}{y - y_0}.$$

Fazendo $y - y_0 = h$,

$$h'(y_0) = \lim_{h\to 0} \frac{f(x_0,y_0 + h) - f(x_0,y_0)}{h} = \frac{\partial f}{\partial y}(x_0,y_0).$$

Novamente, considerando todos os pontos onde é possível calcular $\dfrac{\partial f}{\partial y}(x,y)$, podemos definir a função

$$\begin{aligned} \frac{\partial f}{\partial y} &: D_2 \subset A \subset \mathbb{R}^2 \to \mathbb{R} \\ (x,y) &\mapsto \frac{\partial f}{\partial y}(x,y) = f_y(x,y) \end{aligned}$$

que é chamada função derivada parcial de primeira ordem de f, com respeito à segunda variável.

A reta tangente, neste caso, será denitada por s, veja Figura 13.

Exemplo 11. A derivada parcial em relação a variável y da função$f(x,y) = x^3 + x^2y^3 - 2y^2$ é dada por:

$$\begin{aligned} \frac{\partial f}{\partial y}(x_0,y_0) &= \lim_{y \to y_0} \frac{f(x_0,y) - f(x_0,y_0)}{y - y_0} \\ &= \lim_{y \to y_0} \frac{(x_0^3 + x_0^2 y^3 - 2y^2) - (x_0^3 + x_0^2 y_0^3 - 2y_0^2)}{y - y_0} \\ &= \lim_{y \to y_0} \frac{x_0^2(y^3 - y_0^3) - 2(y^2 - y_0^2)}{y - y_0} \\ &= \lim_{y \to y_0} \frac{x_0^2(y - y_0)(y^2 + yy_0 + y_0^2) - 2(y - y_0)(y + y_0)}{y - y_0} \\ &= \lim_{y \to y_0} x_0^2(y^2 + yy_0 + y_0^2) - 2(y + y_0) \\ &= x_0^2(y_0^2 + y_0^2 + y_0^2) - 2(y_0 + y_0) \\ &= 3x_0^2 y_0^2 - 4y_0. \end{aligned}$$

Exemplo 12. Considerando a função $f(x,y) = 4 - x^2 - y^2$, temos que:

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0,y_0) &= \lim_{x \to x_0} \frac{f(x,y_0) - f(x_0,y_0)}{x - x_0} \\ &= \lim_{x \to x_0} \frac{(4 - x^2 - y_0^2) - (4 - x_0^2 - y_0^2)}{x - x_0} \\ &= \lim_{x \to x_0} \frac{4 - x^2 - y_0^2 - 4 + x_0^2 + y_0^2}{x - x_0} \\ &= \lim_{x \to x_0} \frac{-x^2 + x_0^2}{x - x_0} \\ &= \lim_{x \to x_0} \frac{-(x - x_0)(x + x_0)}{x - x_0} \\ &= \lim_{x \to x_0} -(x + x_0) \\ &= -(x_0 + x_0) \end{aligned}$$

$$= -2x_0.$$

Assim, $f_x(1,1)$ = -2.1 = -2, que representa a taxa de variação de f no ponto $(1,1)$ na direção do eixo x.

Da mesma forma,

$$\begin{aligned}
\frac{\partial f}{\partial y}(x_0,y_0) &= \lim_{y\to y_0} \frac{f(x_0,y)-f(x_0,y_0)}{y-y_0} \\
&= \lim_{y\to y_0} \frac{(4-x_0^2-y^2)-(4-x_0^2-y_0^2)}{y-y_0} \\
&= \lim_{y\to y_0} \frac{4-x_0^2-y^2-4+x_0^2+y_0^2}{y-y_0} \\
&= \lim_{y\to y_0} \frac{-y^2+y_0^2}{y-y_0} \\
&= \lim_{y\to y_0} \frac{-(y-y_0)(y+y_0)}{y-y_0} \\
&= \lim_{y\to y_0} -(y+y_0) \\
&= -(y_0+y_0) \\
&= -2y_0.
\end{aligned}$$

Assim, $f_y(1,1) = -2.1 = -2$, que representa a taxa de variação de f no ponto $(1,1)$ na direção do eixo y.

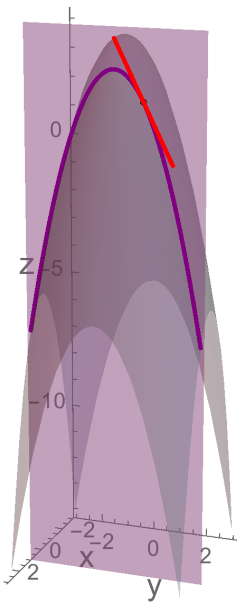

Figura 12 – Reta r, tangente à curva obtida pela intersecção do gráfico de f com o plano yz.

Voltamos agora à ideia de plano tangente como uma “boa aproximação”para uma função de duas variáveis.

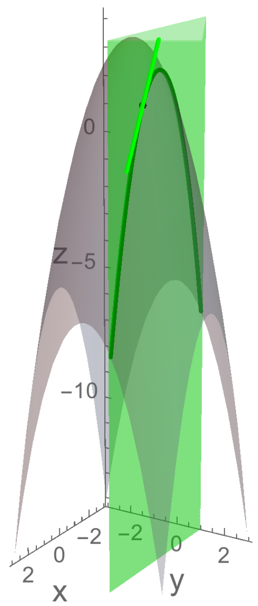

Figura 13 – Reta s, tangente à curva obtida pela intersecção do gráfico de f com o plano xz.

Vamos raciocinar intuitivamente. Seja $z = f(x,y)$ uma superfície que é o gráfico de uma função de duas variáveis.

Pelas considerações sobre derivadas parciais, o plano tangente deve conter as retas r e s com inclinação $\dfrac{\partial f}{\partial y}(x_0,y_0)$ e $\dfrac{\partial f}{\partial x}(x_0,y_0)$, respectivamente. As Figuras 12 e 13 ilustram as retas r e s, respectivamente, juntamente com o gráfico de f, o plano tangente e a curva de intersecção.

Se o plano é dado por :

$$A(x-x_0)+B(y-y_0)+C(z-z_0)=0,$$

ou ainda,

$$(z-z_0)=a(x-x_0)+b(y-y_0), \text{ teremos:}$$

fazendo $x = x_0$ na equação,

$$z-z_0=\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) \qquad \text{(reta r)}$$

e fazendo $y = y_0$ na equação,

$$z-z_0=\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) \qquad \text{(reta s)}$$

$$\text{Portanto, } z-z_0=\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0).$$

Assim, considerando $T(x,y) = f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$, se o plano tangente existir é natural esperarmos que ele tenha esta equação, ou seja, ele será o gráfico da função $T\colon \mathbb{R}^2 \to \mathbb{R}$. A Figura 14 destaca as retas r e s tangentes à superfície enquanto que a Figura 15 exibe o plano tangente à superfície no ponto $((x_0,y_0), f(x_0,y_0))$.

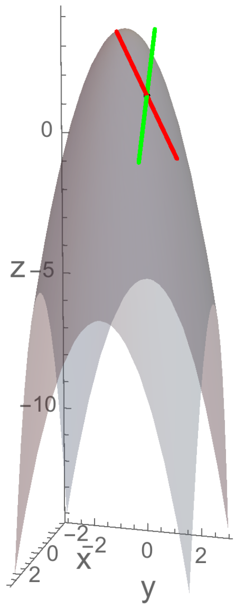

Figura 14 – Retas r e s.

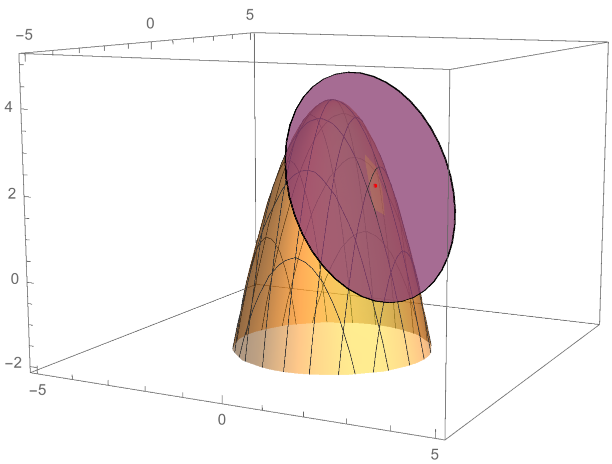

Figura 15 – Plano tangente à superfície.

Façamos algumas considerações. Será que apenas existir as derivadas parciais é suficiente para garantir que a função T seja uma boa aproximação para a função f, no sentido pedido anteriormente? Também, será que funções que possuem derivadas parciais em um ponto (x_0,y_0) são contínuas nesse ponto?

Exemplo 13. Se considerarmos a função $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$ é fácil ver que $f_x(0,0) = f_y(0,0) = 0$, mas f não é contínua na origem.

Entretanto, observamos que a função $f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$ possui as derivadas parciais na origem e é contínua na origem. De fato,

I) $$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{x^3}{x^2+y^2} = \lim_{(x,y)\to(0,0)} \underbrace{x}_{\textit{tende para } 0} \cdot \overbrace{\frac{x^2}{x^2+y^2}}^{\textit{limitada}} = 0 = f(0,0)$$

II) $$f_x(0,0) = \lim_{h\to 0} \frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h^3}{h^2}-0}{h} = 1 \text{ e}$$

$$f_y(0,0) = \lim_{k\to 0} \frac{f(0,0+k)-f(0,0)}{k} = \lim_{k\to 0} \frac{0}{0+k^2} = 0$$

Porém, $f(0,5;0,5) = 0,25$ e, como $T(x,y) = x$, então $T(0,5;0,5) = 0,5$, apesar de $0,5$ ser um valor próximo de $0,25$ ele representa seu dobro, não sendo assim uma boa aproximação.

Logo, apenas a existência das derivadas parciais não implica no conceito de diferenciabilidade que queremos apresentar. Analisando, novamente, o caso para uma variável real é natural pedir que "$f(x,y) - T(x,y) \to 0$ mais rápido do que $(x,y) \to (0,0)$".

Temos, assim, a seguinte definição:

Definição 1.1. Sejam $f : A \to \mathbb{R}$, A é um conjunto aberto de $\mathbb{R}^2$, e $(x_0,y_0) \in A$. Dizemos que f é diferenciável em (x_0,y_0) quando existirem reais a e b, tais que

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0+h,y_0+k) - f(x_0,y_0) - ah - bk}{\|(h,k)\|} = 0.$$

Dizemos que uma função é diferenciável se for diferenciável em todo ponto de seu domínio.

Veremos que, neste caso, $a = \dfrac{\partial f}{\partial x}(x_0,y_0)$ e $b = \dfrac{\partial f}{\partial y}(x_0,y_0)$ pelo seguinte teorema.

Teorema 1. Seja $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ uma função, A um conjunto aberto de $\mathbb{R}^2$, e seja $(x_0,y_0) \in A$. Se f for diferenciável em (x_0,y_0), então f admite suas derivadas parciais neste ponto.

Demonstração. Seja f diferenciável em (x_0,y_0), então existem a e b tais que

$$\lim_{(h,k)\to(0,0)} \frac{E(h,k)}{\|(h,k)\|} = 0, \text{ onde}$$

$$E(h,k) = f(x_0+h, y_0+k) - f(x_0,y_0) - ah - bk, \text{ assim}$$

Se $h < 0$,

$$\lim_{(h,k)\to(0,0)} \frac{E(h,0)}{\|(h,0)\|} = \lim_{h\to 0^-} \left(\frac{f(x_0+h,y_0) - f(x_0,y_0) - ah}{-h} \right) = 0, \text{ então}$$

$$\lim_{h\to 0^-} \left(\frac{-f(x_0+h,y_0) + f(x_0,y_0)}{h} \right) + \lim_{h\to 0^-} \frac{ah}{h} = 0,$$

$$-\lim_{h\to 0^-} \left(\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} \right) = -\lim_{h\to 0^-} \frac{ah}{h},$$

$$\frac{\partial f}{\partial x}(x_0,y_0) = \lim_{h\to 0^-} \frac{ah}{h},$$

$$\frac{\partial f}{\partial x}(x_0,y_0) = a.$$

Se $h > 0$,

$$\lim_{(h,k)\to(0,0)} \frac{E(h,0)}{\|(h,0)\|} = \lim_{h\to 0^+} \left(\frac{f(x_0+h,y_0) - f(x_0,y_0) + ah}{h} \right) = 0, \text{ então}$$

$$\lim_{h\to 0^+} \left(\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} \right) - \lim_{h\to 0^+} \frac{ah}{h} = 0,$$

$$\lim_{h\to 0^+} \left(\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h} \right) = \lim_{h\to 0^+} \frac{ah}{h},$$

$$\frac{\partial f}{\partial x}(x_0,y_0) = a.$$

Da mesma forma temos:

Se $k < 0$,

$$\lim_{(h,k)\to(0,0)} \frac{E(0,k)}{\|(0,k)\|} = \lim_{k\to 0^-} \left(\frac{f(x_0,y_0+k) - f(x_0,y_0) - bk}{-k} \right) = 0 \text{ ,então}$$

$$\lim_{k\to 0^-} \left(\frac{-f(x_0,y_0+k) + f(x_0,y_0)}{k} \right) + \lim_{k\to 0^-} \frac{bk}{k} = 0,$$

$$-\lim_{k\to 0^-} \left(\frac{f(x_0,y_0+k) - f(x_0,y_0)}{k} \right) = -\lim_{k\to 0^-} \frac{bk}{k},$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = b.$$

Se $k > 0$,

$$\lim_{(h,k)\to(0,0)} \frac{E(0,k)}{\|(0,k)\|} = \lim_{k\to 0^+} \left(\frac{f(x_0, y_0 + k) - f(x_0, y_0) - bk}{k} \right) = 0 \text{ ,então}$$

$$\lim_{k\to 0^+} \left(\frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} \right) - \lim_{k\to 0^+} \frac{bk}{k} = 0,$$

$$\lim_{k\to 0^+} \left(\frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} \right) = \lim_{k\to 0^+} \frac{bk}{k},$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = b.$$

Assim provamos que se f for diferenciável em (x_0, y_0), então $a = \frac{\partial f}{\partial x}(x_0, y_0)$ e $b = \frac{\partial f}{\partial y}(x_0, y_0)$ serão os únicos reais para os quais o limite acima é zero.

□

Para uma função com n variáveis o procedimento é o mesmo.

Observações:

1) Para se provar que uma função é diferenciável em (x_0, y_0) é suficiente provar que f admite derivadas parciais em (x_0, y_0) e que

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0 + h, y_0 + k) - f(x_0, y_0) - \frac{\partial f}{\partial x}(x_0, y_0)h - \frac{\partial f}{\partial y}(x_0, y_0)k}{\|(h,k)\|} = 0.$$

2) Se uma das derivadas parciais não existir em (x_0, y_0), então f não será diferenciável nesse ponto.

3) Se ambas as derivadas parciais existirem, mas se o limite acima não for zero, então f não será diferenciável em (x_0, y_0).

4) Se f é diferenciável no ponto (x_0, y_0), o plano $z - f(x_0, y_0) = \frac{\partial f}{\partial x} f(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y} f(x_0, y_0)(y - y_0)$ denomina-se plano tangente ao gráfico de f no ponto $(x_0, y_0, f(x_0, y_0))$.

Exemplo 14. Provaremos que a função $f(x,y) = xy$ é diferenciável. De fato,

$$E(h,k) = f(x_0+h,y_0+k) - f(x_0,y_0) - \frac{\partial f}{\partial x}(x_0,y_0)h - \frac{\partial f}{\partial y}(x_0,y_0)k = hk.$$

Então, $\lim\limits_{(h,k)\to(0,0)} \dfrac{E(h,k)}{\|(h,k)\|} = \lim\limits_{(h,k)\to(0,0)} \dfrac{hk}{\sqrt{h^2+k^2}} = 0.$

Logo, $f(x,y)$ é diferenciável em todo $(x,y) \in \mathbb{R}^2$, ou seja, $f(x,y) = xy$ é uma função diferenciável.

Exemplo 15. Se $f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$, então f não é diferenciável em $(0,0)$. De fato,

$$\lim_{(h,k)\to(0,0)} \frac{E(h,k)}{\|(h,k)\|} =$$

$$\lim_{(h,k)\to(0,0)} \left[(0+h,0+k) - f(0,0) - \frac{\partial f}{\partial x}(0,0)h - \frac{\partial f}{\partial y}(0,0)k\right] . \frac{1}{\sqrt{h^2+k^2}} =$$

$$\lim_{(h,k)\to(0,0)} \frac{\frac{h^2k}{h^2+k^2}}{\sqrt{h^2+k^2}} =$$

$$\lim_{(h,k)\to(0,0)} \frac{h^2k}{(h^2+k^2)\sqrt{h^2+k^2}}.$$

$$G(h,k) = \frac{h^2k}{(h^2+k^2)^{\frac{3}{2}}}.$$

Sendo assim, $\lim\limits_{k\to0} G(0,k) = 0$ e $\lim\limits_{t\to0^+} G(t,t) = \lim\limits_{t\to0^+} \dfrac{t^2t}{2t^2\sqrt{2t^2}} = \lim\limits_{t\to0^+} \dfrac{t}{2t\sqrt{2}} = \dfrac{1}{2\sqrt{2}}.$

Portanto, f não é diferenciável em $(0,0)$.

Exemplo 16. Observamos que a função $f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$ mesmo possuindo as derivadas parcias não possui $T(x,y) = x$ como uma boa aproximação para $f(x,y)$ próximo a $(0,0)$. Mostraremos porque isso ocorre.

Como $f(x,y) - [f(0,0) + f_x(0,0)x + f_y(0,0)y] = E(x,y)$, então

$$\lim_{(x,y)\to(0,0)} \frac{E(x,y)}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{\frac{x^3}{x^2+y^2} - x}{\sqrt{x^2+y^2}} = \lim_{(x,y)\to(0,0)} \frac{-xy^2}{\sqrt{x^2+y^2}(x^2+y^2)}.$$

Se $x = 0$

$$\lim_{(x,y)\to(0,0)} \frac{-xy^2}{\sqrt{x^2+y^2}(x^2+y^2)} = 0.$$

Se $x = y$

$$\lim_{y\to 0} \frac{-y^3}{\sqrt{2y^2}(2y^2)} = \begin{cases} \dfrac{-1}{2\sqrt{2}}, & se\ y > 0 \\ \\ \dfrac{1}{2\sqrt{2}}, & se\ y < 0 \end{cases},$$

o que implica que não existe limite.

Logo, f não é diferenciável em $(0,0)$.

Um outro resultado que gostaríamos de obter é dado pelo seguinte teorema, que relaciona diferenciabilidade e continuidade.

Teorema 2. Se f for diferenciável em (x_0,y_0), então f será contínua em (x_0,y_0).

Demonstração. Seja $f(x,y)$ diferenciável em (x_0,y_0), então existem reais a e b tais que

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0+h,y_0+k)-f(x_0,y_0)-ah-bk}{\|(h,k)\|} = 0, \text{ com } \lim_{(h,k)\to(0,0)} \frac{E(h,k)}{\|(h,k)\|} = 0.$$

Como $\displaystyle\lim_{(h,k)\to(0,0)} ah+bk = 0$, e $\lim_{(h,k)\to(0,0)} E(h,k) = \lim_{(h,k)\to(0,0)} \|(h,k)\| \dfrac{E(h,k)}{\|(h,k)\|} = 0$, temos

$$\lim_{(h,k)\to(0,0)} f(x_0+h,y_0+k) = \lim_{(h,k)\to(0,0)} f(x_0,y_0)+ah+bk+E(h,k).$$

Assim sendo, como $\displaystyle\lim_{(h,k)\to(0,0)} f(x_0,y_0)$ é constante, pois não depende de h e k, $\displaystyle\lim_{(h,k)\to(0,0)} f(x_0+h,y_0+k) = f(x_0,y_0)$.

Portanto, f é contínua em (x_0,y_0).

□

Deste teorema podemos observar que, se f não for contínua em (x_0,y_0), então não será diferenciável em (x_0,y_0).

Exemplo 17. A função $f(x,y) = \begin{cases} \dfrac{2xy^2}{x^2+y^4}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$ não é diferenciável em $(0,0)$. De fato,

$$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{(x,y)\to(0,0)} \frac{2xy^2}{x^2+y^4} = \begin{cases} 1, & \text{se } x = y^2, \\ 0, & \text{se } x = 0. \end{cases}$$

Como f não é contínua em $(0,0)$, não é diferenciável em $(0,0)$.

Quando calculamos as derivadas parciais de uma função $z = f(x,y)$ estamos interessados em calcular a taxa de variação da função nas direções dos eixos x e y .

E se quisermos encontrar a taxa de variação da função numa direção qualquer? Para isso precisamos do conceito de derivada direcional.

Teorema 3. Sejam f uma função de duas variáveis, (x_0,y_0) um ponto do domínio de f e $\vec{u} = (a,b)$ um vetor unitário. Suponhamos que exista $r > 0$ tal que, para $|t| < r$, os pontos da reta $(x,y) = (x_0 + at, y_0 + bt)$ pertencem ao domínio de f. Como $\vec{u} = (a,b)$ é um vetor unitário, a distância de $(x_0 + at, y_0 + bt)$ a (x_0,y_0) é $|t|$.

A taxa média de variação de f, na direção $\vec{u} = (a,b)$, entre os pontos (x_0,y_0) e $(x_0 + at, y_0 + bt)$ é dada por

$$\frac{f(x_0 + at, y_0 + bt) - f(x_0,y_0)}{t}.$$

A derivada direcional de f no ponto (x_0,y_0) e na direção do vetor $\vec{u} = (a,b)$ unitário é

$$\frac{\partial f}{\partial \vec{u}}(x_0,y_0) = \lim_{t\to 0} \frac{f(x_0 + at, y_0 + bt) - f(x_0,y_0)}{t},$$

quando o limite existir.

Observação: Em uma definição mais geral, o conceito de derivada direcional não exige que $\vec{u}$ seja unitário, veja [3].

Exemplo 18. Observamos que a função $f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$ tem derivada direcional no ponto (0,0) em todas as direções. De fato, considerando

$$\overrightarrow{u} = (a,b), \qquad \|\overrightarrow{u}\| = \sqrt{a^2+b^2} = 1$$

$$\begin{aligned} \frac{\partial f}{\partial \overrightarrow{u}}(0,0) &= \lim_{t\to 0} \frac{f(at,bt) - f(0,0)}{t} \\ &= \lim_{t\to 0} \left(\frac{a^3t^3}{a^2t^2 + b^2t^2} \right) \frac{1}{t} \\ &= \lim_{t\to 0} \left(\frac{a^3}{a^2+b^2} \right) \\ &= \frac{a^3}{a^2+b^2} \\ &= a^3. \end{aligned}$$

Portanto, f tem derivada direcional no ponto $(0,0)$ em toda direção, mas f não é diferenciável em $(0,0)$, como já foi mostrado no Exemplo 13.

Logo, ter as derivadas direcionais em todas as direções em um dado ponto (x_0,y_0), não implica em diferenciabilidade .

Uma condição suficiente para garantir a diferenciabilidade de uma função é dada pelo seguinte teorema (ver [1], p. 195).

Teorema 4. Sejam $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, A aberto e $(x_0,y_0) \in A$. Se as derivadas parciais $\dfrac{\partial f}{\partial x}$ e $\dfrac{\partial f}{\partial y}$ existirem em A e forem contínuas no ponto (x_0,y_0), então f será diferenciável neste ponto.

Destacamos a seguir, alguns resultados para funções de várias variáveis a valores reais, que serão utilizados no restante do trabalho.

1.2.3 Diferencial e Gradiente

Seja $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ diferenciável em (x_0,y_0) e consideremos a transformação linear $L : \mathbb{R}^2 \to \mathbb{R}$ dada por

$$L(h,k) = \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k \ .$$

Segue, do que foi visto anteriormente, que $L(h,k)$ é a única transformação linear de $\mathbb{R}^2$ em $\mathbb{R}$ que aproxima o acréscimo $f(x_0+h,y_0+k) - f(x_0,y_0)$ com erro $E(h,k)$ que tende a zero mais rápido que $||(h,k)||$, quando (h,k) tende a $(0,0)$. Isto é,

$$f(x_0+h,y_0+k) - f(x_0,y_0) = \underbrace{\frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k}_{L(h,k)} + E(h,k). \tag{1.1}$$

com $\displaystyle\lim_{(h,k)\to(0,0)} \frac{E(h,k)}{||(h,k)||} = 0.$

A transformação linear L denomina-se **diferencial** de f em (x_0,y_0).

O vetor $\nabla f(x_0,y_0) = \left(\dfrac{\partial f}{\partial x}(x_0,y_0), \dfrac{\partial f}{\partial y}(x_0,y_0)\right)$ denomina-se **vetor gradiente** de f em (x_0,y_0) e pode ser representado pela matriz coluna $\begin{bmatrix} \frac{\partial f}{\partial x}(x_0,y_0) \\ \frac{\partial f}{\partial y}(x_0,y_0) \end{bmatrix}$. Geometricamente, o gradiente é representado como um vetor aplicado em (x_0,y_0). Desta forma, $L(h,k) = \nabla f(x_0,y_0) \cdot (h,k)$, onde o ponto indica o produto escalar entre os vetores.

Seja $T(x,y) = f(x_0,y_0) + \dfrac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \dfrac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$. Sabemos que o gráfico de T é o plano tangente ao gráfico de f no ponto $(x_0,y_0,f(x_0,y_0))$. Fazendo $x = x_0 + h$ e $y = y_0 + k$, tem-se :

$$\underbrace{T(x_0+h,y_0+k) - f(x_0,y_0)}_{T(x_0+h,y_0+k)-T(x_0,y_0)} = \underbrace{\frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k}_{L(h,k)}. \tag{1.2}$$

Dessa forma, $L(h,k)$ é a variação que T sofre quando se passa do ponto (x_0,y_0) ao ponto (x_0+h,y_0+k). Por outro lado, $f(x_0+h,y_0+k) - f(x_0,y_0)$ é a variação em f, quando se passa de (x_0,y_0) a (x_0+h,y_0+k). Logo, por (1.1) e (1.2), temos

$$f(x_0+h,y_0+k) - f(x_0,y_0) \cong \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k,$$

quanto menores forem os módulos de h e k, melhor a aproximação.

Consideremos, agora, a função diferenciável $z = f(x,y)$. Em notação clássica, a diferencial de f em (x,y), relativa aos acréscimos dx e dy é indicada por dz (ou df):

$$dz = \frac{\partial f}{\partial x}(x,y)dx + \frac{\partial f}{\partial y}(x,y)dy.$$

No que segue, iremos nos referir a dz simplesmente como a diferencial de $z = f(x,y)$. Usaremos o símbolo Δz para representar a variação em f, quando se passa de (x,y) a $(x+dx,y+dy)$.

Logo,

$$\Delta z \cong df.$$

Exemplo 19. Consideremos uma caixa, de forma cilíndrica, com tampa, com raio da base igual a 2 cm e altura igual a 5cm, sendo o custo do material usado em sua confecção igual a $0,81$ reais por cm^2. Se as dimensões sofrerem um acréscimo de 1% no raio e 2% na altura vamos determinar o valor aproximado do acréscimo no custo da caixa e o valor exato desse acréscimo.

Como $dr = 1\%$ de 2 é igual a $0,2$ e $dh = 2\%$ de 5 é igual a $0,1$, temos:

$A(r,h) = 2\pi r^2 + 2\pi rh$, assim,

$d_A = (4\pi r + 2\pi h)d_r + 2\pi r d_h = 4\pi.$

Sendo o custo dado por $C(r,h) = 0,81.$área, então $d_C(2,5) = 0,81.4\pi \cong 10,17$ (valor aproximado).

$C(2,5) = 0,81.(2\pi.2^2 + 2\pi.25) \cong 71,21$ (custo sem variação) e $C(2,2;5,1) = 81,69$, logo $\Delta C \equiv 81,69 - 71,21 = 10,48$ (valor exato).

Então, $\dfrac{10,17}{71,21} = 14,28\%$ e $\dfrac{10,48}{71,21} = 14,72\%$, com um erro de $0,44\%$.

1.2.4 Regra da Cadeia

Teorema 5. Sejam $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, A aberto e $\vec{\gamma} : I \to \mathbb{R}$, tais que $\vec{\gamma}(t) \in A$ para todo $t \in I$. Nessas condições, se $\vec{\gamma}$ for diferenciável em t_0, e $(x_0, y_0) = \vec{\gamma}(t_0)$, então a composta $F(t) = f(\vec{\gamma}(t))$ também será diferenciável em t_0 e vale a regra da cadeia:

$$F'(t) = \nabla f(\vec{\gamma}(t)).\vec{\gamma}'(t).$$

Para demonstrar isso, utilizaremos o seguinte lema.

Lema 1. Se $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, com A aberto, for diferenciável em $(x_0, y_0) \in A$, então existirá uma função ϕ definida em A tal que $f(x,y) - f(x_0,y_0) = \nabla f(x_0,y_0)((x,y) - (x_0,y_0)) + \phi(x,y)\|(x,y) - (x_0,y_0)\|$; com $\lim\limits_{(x,y)\to(x_0,y_0)} \phi(x,y) = 0 = \phi(x_0,y_0)$.

Demonstração. Sendo f diferenciável em (x_0,y_0), tem-se:

$$f(x,y) - f(x_0,y_0) = \nabla f(x_0,y_0)((x,y) - (x_0,y_0)) + E(x,y),$$

com $\lim\limits_{(x,y)\to(x_0,y_0)} \dfrac{E(x,y)}{\|(x,y) - (x_0,y_0)\|} = 0.$

Considerando-se

$$\phi(x,y) = \begin{cases} \dfrac{E(x,y)}{\|(x,y) - (x_0,y_0)\|}, & \text{se} \quad (x,y) \neq (x_0,y_0) \\ 0, & \text{se} \quad (x,y) = (x_0,y_0) \end{cases},$$

segue que ϕ é contínua em (x_0,y_0).

□

Pelo lema, substituindo $(x,y) \in A$ por $\vec{\gamma}(t)$ e (x_0,y_0) por $\vec{\gamma}(t_0)$, temos:

$$f(\vec{\gamma}(t)) - f(\vec{\gamma}(t_0)) = \nabla f(\vec{\gamma}(t_0))(\vec{\gamma}(t) - \vec{\gamma}(t_0)) + \phi(\vec{\gamma}(t))\|\vec{\gamma}(t) - \vec{\gamma}(t_0)\|,$$

e dividindo-se todos os termos por $t - t_0$, $t \neq t_0$, segue que:

$$\frac{f(\vec{\gamma}(t))-f(\vec{\gamma}(t_0))}{t-t_0}=\frac{\nabla f(\vec{\gamma}(t_0))(\vec{\gamma}(t)-\vec{\gamma}(t_0))}{t-t_0}+\frac{\phi(\vec{\gamma}(t))\|\vec{\gamma}(t)-\vec{\gamma}(t_0)\|}{t-t_0},$$

como

$$\frac{\|\vec{\gamma}(t)-\vec{\gamma}(t_0)\|}{t-t_0}=\frac{|t-t_0|}{t-t_0}\left\|\frac{\vec{\gamma}(t)-\vec{\gamma}(t_0)}{t-t_0}\right\|,$$

$$\lim_{t\to t_0}\phi(\vec{\gamma}(t))\frac{|t-t_0|}{t-t_0}=0,\text{ pois }\frac{|t-t_0|}{t-t_0}\text{ é limitada.}$$

Como

$$\lim_{t\to t_0}\left\|\frac{\vec{\gamma}(t)-\vec{\gamma}(t_0)}{t-t_0}\right\|=\|\vec{\gamma}'(t_0)\|,\text{ temos então:}$$

$$\lim_{t\to t_0}\phi(\vec{\gamma}(t))\frac{\|\vec{\gamma}(t)-\vec{\gamma}(t_0)\|}{t-t_0}=0.$$

Então,

$$\lim_{t\to t_0}\frac{f(\vec{\gamma}(t))-f(\vec{\gamma}(t_0))}{t-t_0}=\nabla f(\vec{\gamma}(t_0))\vec{\gamma}'(t_0)+0$$

Portanto,

$$F'(t_0)=\nabla f(\vec{\gamma}(t_0)).\vec{\gamma}'(t_0).$$

□

Exemplo 20. Considerando a função $z = x^2y + 3xy^4$ e sendo $x = \operatorname{sen}(2t)$ e $y = \cos t$ vamos determinar $\frac{dz}{dt}$.

Calculando as derivadas parcias em relação a x e a y temos: $z_x = 2xy + 3y^4$ e $z_y = x^2 + 12xy^3$.

Pela Regra da Cadeia,

$\frac{dz}{dt} = \frac{\partial z}{\partial x}.\frac{dx}{dt} + \frac{\partial z}{\partial y}.\frac{dy}{dt}$

$= (2sen(2t).cos(t)+3cos^4(t)).(2cos(2t)) + (sen^2(2t)+12sen(2t).cos^3(t)).(-sen(t))$

$= 4sen(2t).cos(2t).cos(t)+6cos^4(t).cos(2t)$ - $sen^2(2t).sen(t)-12sen(2t).sen(t).cos^3(t)$

Então, $\frac{dz}{dt}(0) = 6$ é a taxa de crescimento (ou velocidade) com que a função está variando ao longo da curva $(\operatorname{sen} 2t, \cos t)$, passando por $(0,1)$.

1.2.5 Funções Definidas Implicitamente

Uma função $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, sendo A aberto, é de classe C^k em A, se f admitir todas as derivadas parcias em todos os pontos de A e $\dfrac{\partial f}{\partial x}, \dfrac{\partial f}{\partial y} : A \to \mathbb{R}$ forem de classe C^{k-1}, onde k e um inteiro positivo. Dizemos que $f : A \to \mathbb{R}$ é de classe C^0 quando ela for contínua em A. Dessa forma, se f possui todas as derivadas parciais e estas são contínuas, f é de classe C^1.

Teorema 6. (Teorema da Função Implícita) Sejam $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ de classe C^1 e $(x_0, y_0) \in A$, com $f(x_0, y_0) = 0$. Nessas condições, se $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$, então existirão intervalos abertos I e J, com $x_0 \in I$ e $y_0 \in J$, tais que, para cada $x \in I$, existe um único $g(x) \in J$, com $f(x, g(x)) = 0$. A função $g : I \to J$ é diferenciável e $g'(x) = \dfrac{-\dfrac{\partial f}{\partial x}(x, g(x))}{\dfrac{\partial f}{\partial y}(x, g(x))}$.

Demonstração. Suponhamos, sem perda de generalidade, que $\dfrac{\partial f}{\partial y}(x_0, y_0) > 0$. $\dfrac{\partial f}{\partial y}$ é contínua em (x_0, y_0), pois, por hipótese, f é de classe C^1. Como $\dfrac{\partial f}{\partial y}(x_0, y_0) > 0$, pelo Teorema da Conservação do Sinal existe uma bola aberta B de centro (x_0, y_0), que podemos supor contida em A, pois A é aberto, tal que $\dfrac{\partial f}{\partial y}(x, y) > 0$ em B.

Sejam y_1 e y_2 tais que $y_1 < y_0 < y_2$, com (x_0, y_1) e (x_0, y_2) em B. Fixado x_0, consideremos a função $z = f(x_0, y), y \in [y_1, y_2]$. Como $\dfrac{\partial f}{\partial y}(x_0, y) > 0$ para todo $y \in [y_1, y_2]$, segue que z é estritamente crescente em $[y_1, y_2]$. Tendo em vista que $f(x_0, y_0) = 0$, resulta $f(x_0, y_1) < 0$ e $f(x_0, y_2) > 0$. Seja $J =]y_1, y_2[$; observe que $y_0 = g(x_0)$ é o único número em J tal que $f(x_0, y_0) = 0$. Tendo em vista $f(x_0, y_1) < 0$ e $f(x_0, y_2) > 0$ e pela continuidade de f, existe um intervalo aberto I com $x_0 \in I$, onde $I = I_1 \cap I_2$ e I_1 é um intervalo contendo x_0 tal que $f(x, y_1) < 0$ e I_2 é um intervalo contendo x_0 tal que $f(x, y_2) > 0$, garantido pela continuidade de f, tal que para todo $x \in I$, (x, y_1) e (x, y_2) pertencem a B, com $f(x, y_1) < 0$ e $f(x, y_2) > 0$. Como $\dfrac{\partial f}{\partial y}(x, y) > 0$ em B, para todo $x \in I$, a função $z = f(x, y)$, com x fixo, é estritamente crescente em $[y_1, y_2]$; tendo em vista que $f(x, y_1) < 0$ e $f(x, y_2) > 0$, pelo teorema do valor intermediário e pelo fato de $f(x, y)$ ser estritamente crescente em $]y_1, y_2[$, para este x fixo, existirá um único $y \in]y_1, y_2[$ que podemos chamar de $g(x)$, tal que $f(x, y) = f(x, g(x)) = 0$. A função $g : I \to J$ está definida implicitamente pela equação $f(x, y) = 0$.

Observação. Para todos $\bar{y}_1$ e $\bar{y}_2$, com $y_1 < \bar{y}_1 < y_0 < \bar{y}_2 < y_2$, procedendo como acima, encontramos um intervalo aberto $I_1 \subset I$, com $x_0 \in I_1$, tal que $x \in I_1 \Rightarrow g(x) \in]\bar{y}_1, \bar{y}_2[$; logo, g é contínua em x_0. O mesmo raciocínio pode ser feito para mostrar que que g é contínua em I.

Como f é diferenciável em (x_0, y_0), então existem funções $\varphi_1(x,y)$ e $\varphi_2(x,y)$, definidas no domínio de f, tais que

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + \varphi_1(x,y)(x-x_0) + \varphi_2(x,y)(y-y_0)$$

com $\lim_{(x,y)\to(x_0,y_0)} \varphi_1(x,y) = 0 = \varphi_1(x_0,y_0)$ e $\lim_{(x,y)\to(x_0,y_0)} \varphi_2(x,y) = 0 = \varphi_2(x_0,y_0)$.

De fato, pelo Lema 1, temos que

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + \varphi(x,y)||(x,y)-(x_0,y_0)||,$$

onde $\lim_{(x,y)\to(x_0,y_0)} \varphi(x,y) = 0 = \varphi(x_0,y_0)$.

Para $(x,y) \neq (x_0,y_0)$, podemos considerar;

$$\varphi(x,y)||(x,y)-(x_0,y_0)|| = \varphi(x,y)\frac{(x-x_0)^2+(y-y_0)^2}{||(x,y)-(x_0,y_0)||}$$

$$= \varphi(x,y)\frac{x-x_0}{||(x,y)-(x_0,y_0)||}(x-x_0) + \varphi(x,y)\frac{y-y_0}{||(x,y)-(x_0,y_0)||}(y-y_0).$$

Basta tomar

$$\varphi_1(x,y) = \begin{cases} \varphi(x,y)\dfrac{x-x_0}{||(x,y)-(x_0,y_0)||} & \text{se } (x,y) \neq (x_0,y_0), \\ 0 & \text{se } (x,y) = (x_0,y_0), \end{cases}$$

e

$$\varphi_2(x,y) = \begin{cases} \varphi(x,y)\dfrac{y-y_0}{||(x,y)-(x_0,y_0)||} & \text{se } (x,y) \neq (x_0,y_0), \\ 0 & \text{se } (x,y) = (x_0,y_0). \end{cases}$$

Daí, substituindo $y = g(x)$ e $y_0 = g(x_0)$ em $f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + \varphi_1(x,y)(x-x_0) + \varphi_2(x,y)(y-y_0)$ e lembrando que $f(x,g(x)) = 0$ e $f(x_0,g(x_0)) = 0$ resulta, após dividir por $x-x_0$, com $x \neq x_0$, que:

$$0 = f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,g(x_0)) + \frac{\partial f}{\partial y}(x_0,g(x_0))\frac{g(x)-g(x_0)}{x-x_0} + \varphi_1(x,g(x)) +$$

$$+\varphi_2(x,g(x))\frac{g(x)-g(x_0)}{x-x_0}.$$

Pelo fato de g ser contínua em x_0, se x converge para x_0, $g(x)$ converge para $g(x_0)$, e $\frac{\partial f}{\partial y}(x_0,y_0) \neq 0$, resulta em $g'(x_0) = \dfrac{-\frac{\partial f}{\partial x}(x,g(x))}{\frac{\partial f}{\partial y}(x,g(x))}$, se $\frac{\partial f}{\partial y}(x,g(x)) \neq 0$.

$\square$

Observação: Se a hipótese $\frac{\partial f}{\partial y}(x_0,y_0) \neq 0$ for substituída por $\frac{\partial f}{\partial x}(x_0,y_0) \neq 0$, então existirão intervalos abertos I e J com $x_0 \in I$ e $y_0 \in J$, tais que, para cada $y \in J$, existirá um único $h(y) \in I$, com $f(h(y),y) = 0$. a função $h : J \to I$ será diferenciável e

$$h'(y) = \frac{-\dfrac{\partial f}{\partial y}(h(y),y)}{\dfrac{\partial f}{\partial x}(h(y),y)}, \text{ se } \frac{\partial f}{\partial x}(h(y),y) \neq 0.$$

Exemplo 21. Consideremos a seguinte equação: $x^2 + y^2 = 1$.

Esta relação não define y como função de x, nem x como função de y. Entretanto, se considerarmos o conjunto $U_1 = \{(x,y) \in \mathbb{R}^2 / y > 0\}$, $x^2 + y^2 = 1$ então $y = \sqrt{1-x^2}$, $\forall (x,y) \in U_1$ e, portanto, esta relação define implicitamente y como função de x, isto é, $x^2 + (f(x))^2 = 1$, onde $f(x) = \sqrt{1-x^2}$.

Analogamente, para os conjuntos $U_2 = \{(x,y) \in \mathbb{R}^2 / y < 0\}$, $U_3 = \{(x,y) \in \mathbb{R}^2 / x > 0\}$ e $U_4 = \{(x,y) \in \mathbb{R}^2 / x < 0\}$, $x^2 + y^2 = 1$ define localmente y como função de x ou x como função de y.

Exemplo 22. Analisando a função $x^3 + y^3 = 6xy$ (ver Figura 16), y pode ser visto como uma função de x definida implicitamente pela relação em torno do ponto $(3,3)$ garantido pelo Teorema da Função Implícita. De fato, se

$f(x,y) = x^3 + y^3$ - $6xy$, então $f(3,3) = 0$ e $f_x(x,y) = 3x^2$ - $6y$, $f_y(x,y) = 3y^2$ - $6x$ e $f_y(3,3) > 0$.

Assim, $y' = g'(x) = -\dfrac{(3x^2-6y)}{3y^2-6x} = \dfrac{6y-3x^2}{3y^2-6x} = \dfrac{2y-x^2}{y^2-2x}$, desde que y^2 - $2x \neq 0$.

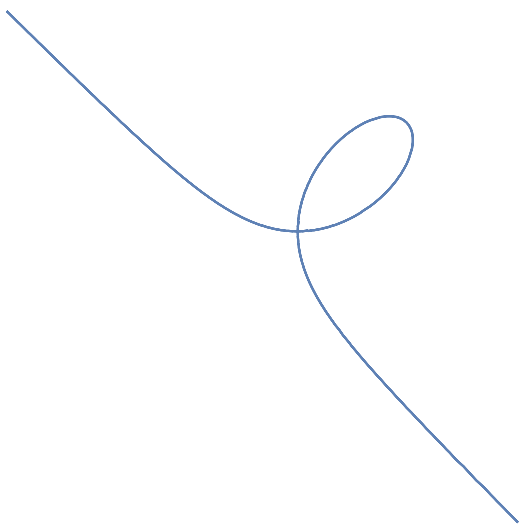

Figura 16 – Fólio de Descartes, curva definida pela equação $x^3 + y^3 = 6xy$.

1.2.6 Derivadas Parciais de Ordem Superior

Seja $f : A \subset \mathbb{R}^2 \to \mathbb{R}$ diferenciável em A e considere suas derivadas parciais

$$\frac{\partial f}{\partial x} : A_1 \subseteq A \subset \mathbb{R}^2 \to \mathbb{R} \text{ e } \frac{\partial f}{\partial y} : A_2 \subseteq A \subset \mathbb{R}^2 \to \mathbb{R}.$$

Se suas derivadas parciais forem diferenciáveis em $(x_0, y_0) \in A_1 \cap A_2$, dizemos que f é duas vezes diferenciável em (x_0, y_0). Neste caso, existem as derivadas parciais de segunda ordem da f, dadas por:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = f_{xx}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = f_{xy}$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = f_{yy}$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = f_{yx}$$

Quando f é duas vezes diferenciável em todos os pontos de A, ficam definidas as funções $\frac{\partial^2 f}{\partial x^2}, \frac{\partial^2 f}{\partial y \partial x}, \frac{\partial^2 f}{\partial y^2}, \frac{\partial^2 f}{\partial x \partial y} : A \to \mathbb{R}$. Se todas essas derivadas parciais forem diferenciáveis num ponto, diremos que f é três vezes diferenciável neste ponto. Obtemos, assim, as derivadas parciais de terceira ordem da f. Por exemplo,

$$\frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial y \partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)\right) = f_{xyx}.$$

Como f é C^1, segue que f é diferenciável, segue que se f é C^k, f é k vezes diferenciável.

Exemplo 23. Considerando a função $f(x,y) = x^3 + x^2y^3$ - $2y^2$, vamos determinar suas derivadas parciais de segunda ordem.

Temos, $f_x(x,y) = 3x^2 + 2xy^3$ e $f_y(x,y) = 3y^2x^2$ - $4y$.

Logo,

$$f_{xx}(x,y) = 6x + 2y^3,\ f_{xy}(x,y) = 6xy^2,\ f_{yy}(x,y) = 6x^2y \text{ - } 4 \text{ e } f_{yx}(x,y) = 6xy^2.$$

Exemplo 24. Seja a função $f(x,y)= x\ cos(y)\ +\ y\ e^x$, determinaremos suas derivadas parciais de segunda ordem.

Temos, $f_x(x,y) = cos(y) + y\ e^x$ e $f_y(x,y) = -x\ sen(y) + e^x$.

Logo,

$$f_{xx}(x,y) = y\ e^x,\ f_{xy}(x,y) = -sen(y) + e^x,$$

$$f_{yx}(x,y) = -sen(y) + e^x \text{ e } f_{yy}(x,y) = -x\ cos(y).$$

Observe que nos exemplos anteriores temos $f_{xy} = f_{yx}$, porém nem sempre isto é válido, o que será verificado no exemplo abaixo.

Exemplo 25. Considerando a função $f(x,y) = \begin{cases} \dfrac{xy^3}{x^2+y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}.$

Calcularemos f_{xy} e f_{yx}.

$$f_x(x,y) = \frac{y^3(x^2+y^2) - xy^3 2x}{(x^2+y^2)^2} = \frac{y^3x^2+y^5-2x^2y^3}{(x^2+y^2)^2}, \text{ para } (x,y) \neq (0,0)$$

$$f_x(0,0) = \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0}{h^2} - 0}{h} = 0$$

Portanto,

$$f_x(x,y) = \begin{cases} \dfrac{y^5 - x^2y^3}{(x^2+y^2)^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases},$$

$$f_y(x,y) = \frac{3xy^2(x^2+y^2) - xy^3 2y}{(x^2+y^2)^2} = \frac{3x^3y^2+xy^4}{(x^2+y^2)^2}, \text{ para } (x,y) \neq (0,0),$$

$$f_y(0,0) = \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = \lim_{h\to 0} 0 = 0.$$

Portanto,

$$f_y(x,y) = \begin{cases} \dfrac{3x^3y^2+xy^4}{(x^2+y^2)^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}.$$

Assim,

$$f_{yx}(0,0) = \lim_{h\to 0} \frac{\left[\frac{\partial f}{\partial y}(0+h,0) - \frac{\partial f}{\partial y}(0,0)\right]}{h} = \lim_{h\to 0} \frac{\frac{0}{h^4} - 0}{h} = 0,$$

$$f_{xy}(0,0) = \lim_{h\to 0} \frac{\left[\frac{\partial f}{\partial x}(0,0+h) - \frac{\partial f}{\partial x}(0,0)\right]}{h} = \lim_{h\to 0} \frac{\frac{h^5}{h^4} - 0}{h} = 1.$$

Portanto, $f_{yx}(0,0) \neq f_{xy}(0,0)$.

Os resultados obtidos nos exemplos acima, podem ser observados no seguinte teorema(para uma demonstração ver [3] p.148):

Teorema 7. Teorema de Schwarz: Seja $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, A aberto. Se f é duas vezes diferenciável em (x,y), então:

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) \ .$$

Como podemos ver facilmente, as derivadas parciais mistas do Exemplo 25 são contínuas para $(x,y) \neq (0,0)$. Logo, pelo teorema de Schwarz, $\dfrac{\partial^2 f}{\partial x \partial y}(x,y) = \dfrac{\partial^2 f}{\partial y \partial x}(x,y)$, para $(x,y) \neq (0,0)$.

1.3 Interpretação geométrica do vetor gradiente

Faremos uma interpretação geométrica do vetor gradiente.

Sejam $f : A \subset \mathbb{R}^2 \to \mathbb{R}$, A aberto em $\mathbb{R}^2$, f de classe C^1 e (x_0,y_0) um ponto da curva de nível $f(x,y) = c$.

Suponhamos que $\nabla f(x_0,y_0) \neq (0,0)$, isto é, $\dfrac{\partial f}{\partial x}(x_0,y_0) \neq 0$ ou $\dfrac{\partial f}{\partial y}(x_0,y_0) \neq 0$. Vamos supor, sem perda de generalidade, que $\dfrac{\partial f}{\partial y}(x_0,y_0) \neq 0$. Pelo Teorema da Função Implícita, existem intervalos abertos $I, J \subset \mathbb{R}$ com $x_0 \in I$ e $y_0 \in J$ e uma função $y = y(x)$, $y : I \to J$ tal que $f(x,y(x)) = c$, com $y(x_0) = y_0$ e $y'(x) = \dfrac{-\dfrac{\partial f}{\partial x}(x,y(x))}{\dfrac{\partial f}{\partial y}(x,y(x))}$, $x \in I$.

Isto é, existe uma curva $\vec{\gamma}$, diferenciável, passando por (x_0,y_0) cuja imagem está contida na curva de nível $f(x,y) = c$, a saber, $\vec{\gamma} : I \subset \mathbb{R} \to \mathbb{R}^2$ tal que $\vec{\gamma}(t) = (t, y(t))$, com $\vec{\gamma}(t_0) = (x_0,y_0)$.

Pela regra da cadeia, considerando f restrita a curva $\vec{\gamma}$, temos

$$\frac{d}{dt}[f(\vec{\gamma}(t))] = \frac{d}{dt}(c)$$

$$\nabla f(\vec{\gamma}(t)).\vec{\gamma}'(t) = 0, \ \forall t \in I.$$

Em particular, $\nabla f(\vec{\gamma}(t_0)).\vec{\gamma}'(t_0) = 0$.

Como $\vec{\gamma}'(t_0)$ é um vetor tangente a curva $\vec{\gamma}$ em t_0, temos que $\nabla f(x_0,y_0)$ é perpendicular a $\vec{\gamma}$, em $\vec{\gamma}(t_0) = (x_0,y_0)$.

Dizemos que $\nabla f(x,y)$ é um vetor normal à curva de nível $f(x,y)=c$, em (x_0,y_0). A reta passando por (x_0,y_0) é perpendicular ao $\nabla f(x_0,y_0)$ e denomina-se reta tangente, em (x_0,y_0), à curva de nível $f(x,y)=c$. A equação desta reta é dada por

$$\nabla f(x_0,y_0)[(x,y)-(x_0,y_0)]=0.$$

Podemos estender este resultado para funções de três ou mais variáveis mostrando que $\nabla f(x_0,y_0,z_0)$ é normal em (x_0,y_0,z_0) à superfície de nível $f(x,y,z)=c$, que passa por (x_0,y_0,z_0).

1.4 Polinômio de Taylor para uma função com 2 variáveis

Polinômio de Taylor de ordem 1

Seja f de classe C^2 no aberto $A\subset\mathbb{R}^2$. Sejam $(x_0,y_0)\in A$ e $(h,k)\neq(0,0)$ tais que o segmento de extremidades (x_0,y_0) e (x_0+h,y_0+k) esteja contido em A. Considere a função g dada por:

$$g(t)=f(\underbrace{x_0+ht}_{x},\underbrace{y_0+kt}_{y}),\, t\in[0,1].$$

A função g fornece os valores que f assume em pontos do segmento de extremidades (x_0,y_0) e (x_0+h,y_0+k).Utilizando a fórmula de Taylor para polinômios de uma variável, temos que :

$$g(1)=g(0)+g'(0)(1-0)+\frac{g''(\bar{t})}{2}(1-0)^2,$$

para algum $\bar{t}\in]0,1[$.

Pela regra da cadeia, considerando $\vec{\gamma}(t)=(x(t),y(t))=(x_0+ht,y_0+kt)$,$t\in[0,1]$ temos,

$$g'(t)=\frac{\partial f}{\partial x}(x_0+ht,y_0+kt)\frac{dx}{dt}+\frac{\partial f}{\partial y}(x_0+ht,y_0+kt)\frac{dy}{dt}$$

$$=\frac{\partial f}{\partial x}(x,y)h+\frac{\partial f}{\partial y}(x,y)k \quad \text{e,}$$

$$g''(t)=\frac{d}{dt}g'(t)=\frac{d}{dt}\left(\frac{\partial f}{\partial x}(x,y)\right)h+\frac{d}{dt}\left(\frac{\partial f}{\partial y}(x,y)\right)k$$

$$=\left(\frac{\partial^2 f}{\partial x^2}(x,y)h+\frac{\partial^2 f}{\partial x\partial y}(x,y)k\right)h+\left(\frac{\partial^2 f}{\partial y\partial x}(x,y)h+\frac{\partial^2 f}{\partial y^2}(x,y)k\right)k$$

$$= \frac{\partial^2 f}{\partial x^2}(x,y)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x,y)hk + \frac{\partial^2 f}{\partial y^2}(x,y)k^2,$$

usando o Teorema de Schwarz. Como,

$$\begin{cases} g(1) &= & f(x_0+h, y_0+k) \\ g(0) &= & f(x_0,y_0) \\ g'(0) &= & \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k \end{cases}$$

definindo $(\bar{x},\bar{y}) = (x_0+h\bar{t}, y_0+k\bar{t})$, temos que $(\bar{x},\bar{y})$ pertence ao segmento de extremidades (x_0,y_0) e (x_0+h, y_0+k), se $\bar{t} \in]0,1[$.

Se $g(1) = g(0) + g'(0)(1-0) + \frac{g''(\bar{t})}{2}(1-0)^2$, para algum $\bar{t} \in (0,1)$, então:

$f(x_0+h, y_0+k)=$

$$f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k + \underbrace{\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y})hk + \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y})k^2\right)}_{E_1(h,k)}.$$

Para $(x,y) = (x_0+h, y_0+k)$,

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) +$$

$$+\frac{1}{2}\left(\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})(x-x_0)^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y})(x-x_0)(y-y_0) + \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y})(y-y_0)^2\right).$$

O polinômio $P_1(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)$ é denominado polinômio de Taylor de grau 1 de f em torno de (x_0,y_0) e $E_1(x,y)$ é o erro que se comete ao aproximar $f(x,y)$ por $P_1(x,y)$. Esta formulação para o erro é denominada erro ou resto de Lagrange.

Polinômio de Taylor de Ordem 2

Seja f de Classe C^3 no aberto $A \subset \mathbb{R}^2$. Sejam $(x_0,y_0) \in A$ e $(h,k) \neq (0,0)$ tais que o segmento de extremidades (x_0,y_0) e (x_0+h, y_0+k) esteja contido em A. Considere a função:

$$g(t) = f(x+ht, y+kt), t \in [0,1].$$

Pelo que foi observado anteriormente,

$$g'(t) = \frac{\partial f}{\partial x}(x,y)h + \frac{\partial f}{\partial y}(x,y)k \text{ e } g''(t) = \frac{\partial^2 f}{\partial x^2}(x,y)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x,y)hk + \frac{\partial^2 f}{\partial y^2}(x,y)k^2.$$

Ainda,

$$g'''(t) = \frac{\partial}{\partial x}\left(\frac{\partial^2 f}{\partial x^2}(x,y)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(x,y)hk + \frac{\partial^2 f}{\partial y^2}(x,y)k^2\right)\frac{dx}{dt} +$$

$+\frac{\partial}{\partial y}\left(\frac{\partial^2 f}{\partial x^2}(x,y)h^2+2\frac{\partial^2 f}{\partial x\partial y}(x,y)hk+\frac{\partial^2 f}{\partial y^2}(x,y)k^2\right)\frac{dy}{dt}$

$=\frac{\partial f^3}{\partial x^3}(x,y)h^3+2\frac{\partial f^3}{\partial x^2\partial y}(x,y)h^2k+\frac{\partial f^3}{\partial x\partial y}(x,y)hk^2+\frac{\partial f^3}{\partial y\partial x^2}(x,y)h^2k+$

$+2\frac{\partial f^3}{\partial x\partial y^2}(x,y)hk^2+\frac{\partial f^3}{\partial y^3}(x,y)k^3$

$=\frac{\partial f^3}{\partial x^3}(x,y)h^3+3\frac{\partial f^3}{\partial x^2\partial y}(x,y)h^2k+3\frac{\partial f^3}{\partial x\partial y^2}(x,y)hk^2+\frac{\partial f^3}{\partial y^3}(x,y)k^3.$

Substituindo no polinômio de segundo grau em uma variável:

$g(1)=g(0)+g'(0)(1-0)+\frac{g"(0)}{2!}(1-0)^2+\frac{g'''(0)}{3!}(1-0)^3$, obtemos:

$f(x,y)=f(x_0,y_0)+\frac{\partial f}{\partial x}(x_0,y_0)(x-x_0)+\frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)+$

$+\frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(x_0,y_0)(x-x_0)^2+2\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)(x-x_0)(y-y_0)+\frac{\partial^2 f}{\partial y^2}(x_0,y_0)(y-y_0)^2\right]+E_2(\bar{x},\bar{y})$

$E_2(\bar{x},\bar{y})=\frac{1}{3!}[\frac{\partial^3 f}{\partial x^3}(\bar{x},\bar{y})(x-x_0)^3+3\frac{\partial^3 f}{\partial x^2\partial y}(\bar{x},\bar{y})(x-x_0)^2(y-y_0)+ +3\frac{\partial^3 f}{\partial x\partial y^2}(\bar{x},\bar{y})(x-$
$x_0)(y-y_0)^2+\frac{\partial^3 f}{\partial y^3}(\bar{x},\bar{y})(y-y_0)^3].$

Utilizando o mesmo procedimento para funções de n variáveis, generalizamos o conceito para um polinômio de Taylor de ordem n :

$$f(x_0+h,y+k)=f(x_0,y_0)+\sum_{n=1}^{n}\frac{1}{n!}\left[\sum_{p=0}^{n}\binom{n}{p}\frac{\partial^n f(x_0,y_0)}{\partial x^{n-p}\partial y^p}h^{n-p}k^p\right]+E(h,k)$$

$$E(h,k)=\frac{1}{(n+1)!}\sum_{p=0}^{n+1}\binom{n+1}{p}\frac{\partial^{n+1}f(\bar{x},\bar{y})}{\partial x^{n+1-p}\partial y^p}h^{n+1-p}k^p,$$

para algum $(\bar{x},\bar{y})$ pertencente ao segmento de extremidades $(x_0,y_0),(x_0+h,y_0+k)$.

□

Exemplo 26. O polinômio de Taylor de ordem 2 para a função $f(x,y)=xseny$, em torno do ponto $(0,0)$ é dado por $P(x,y)=xy$ e o erro que se comete ao fazer essa aproximação é dado pela expressão

$$E(x,y)=\frac{1}{3!}[(-3sen\bar{y})xy^2-(\bar{x}cos\bar{y})y^3].$$

Assim, se considerarmos que $|x|<1$, temos uma estimativa para $E(x,y)$, isto é,

$$|E(x,y)|<\frac{|y|^2}{2}\left[|x|+\frac{|y|}{3}\right].$$

Por exemplo, o valor de $f(0.01,0.01)$ pode ser aproximado por 10^{-4} e o erro que se comete nesta aproximação é menor que $\frac{2}{3}.10^{-6}$.

1.5 Matriz Hessiana

Vamos generalizar o conceito de diferenciabilidade para funções de $\mathbb{R}^m$ em $\mathbb{R}^n$.

Seja $A \subset \mathbb{R}^m$ um conjunto aberto. Dizemos que uma aplicação f: $A \to \mathbb{R}^n$ é diferenciável em um ponto $a \in A$ quando existir uma transformação linear $T : \mathbb{R}^m \to \mathbb{R}^n$ tal que $f(a+v) - f(a) = T.v + r(v)$ onde $\lim\limits_{v \to 0} \dfrac{r(v)}{||v||}=0$.

T é chamada de derivada de f no ponto a e indicada por $f'(a)$ ou $Df(a)$ e $f'(a) \in \mathscr{L}(\mathbb{R}^m,\mathbb{R}^n)$, o conjunto das transformações lineares de $\mathbb{R}^m$ em $\mathbb{R}^n$.

Se $f(x) = (f_1(x), f_2(x), f_3(x), ..., f_n(x))$, onde cada função coordenada $f_i(x) : A \subset \mathbb{R}^m \to \mathbb{R}$, a aplicação $f : A \subset \mathbb{R}^m \to \mathbb{R}^n$ é diferenciável em $a \in A$ se, e somente se, cada uma das funções coordenadas f_1, ..., $f_n : A \to \mathbb{R}$ é diferenciável nesse ponto. Pelo que foi visto anteriormente, se $f_i(x) : A \subset \mathbb{R}^m \to \mathbb{R}$, então $f_i'(a).v = df_i(a).v$, onde $df_i(a)$ é o diferencial de f_i em a relativa ao acréscimo v. Assim,

$$f'(a).v = (f_1'(a).v, f_2'(a).v, ..., f_n'(a).v) = (df_1(a).v, df_2(a).v, ..., df_n(a).v).$$

A matriz dessa transformação linear, relativa às bases canônicas do $\mathbb{R}^m$ e do $\mathbb{R}^n$ é dada por:

$$\mathrm{J}(f)(a) = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1}(a) & \dfrac{\partial f_1}{\partial x_2}(a) & \cdots & \dfrac{\partial f_1}{\partial x_m}(a) \\ \dfrac{\partial f_2}{\partial x_1}(a) & \dfrac{\partial f_2}{\partial x_2}(a) & \cdots & \dfrac{\partial f_2}{\partial x_m}(a) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial f_n}{\partial x_1}(a) & \dfrac{\partial f_n}{\partial x_2}(a) & \cdots & \dfrac{\partial f_n}{\partial x_m}(a) \end{bmatrix}$$

Tal matriz é chamada de matriz jacobiana de f no ponto a. Obviamente, se $f : A \subset \mathbb{R}^n \to \mathbb{R}$, $\mathrm{J}(f)(a)$ é dada por :

$$\mathrm{J}(f)(a) = \begin{bmatrix} \dfrac{\partial f}{\partial x_1}(a) & \dfrac{\partial f}{\partial x_2}(a) & \cdots & \dfrac{\partial f}{\partial x_n}(a) \end{bmatrix} = \nabla f(a).$$

Para cada $v = (\alpha_1, \ldots, \alpha_n)$, escrevemos $d^2 f(a) \cdot v^2 = \sum_{i,j} \dfrac{\partial^2 f}{\partial x_i \partial x_j}(a)\alpha_i\alpha_j$. A forma $d^2 f(a) : \mathbb{R}^n \to \mathbb{R}$ chama-se a segunda diferencial da funcão f no ponto a ou a forma Hessiana da funcão f no ponto a. Tal transformação é uma forma quadrática, objeto que estudaremos no próximo capítulo. A forma hessiana de uma função duas vezes diferenciável será indicada por $\mathscr{H}(f)(a)$ e sua representação matricial é dada pela matriz

$$\mathscr{H}(f)(a) = \begin{bmatrix} \dfrac{\partial^2 f_1}{\partial x_1^2}(a) & \dfrac{\partial^2 f_1}{\partial x_2 \partial x_1}(a) & \cdots & \dfrac{\partial^2 f_1}{\partial x_n \partial x_1}(a) \\ \dfrac{\partial^2 f_2}{\partial x_1 \partial x_2}(a) & \dfrac{\partial^2 f_2}{\partial x_2^2}(a) & \cdots & \dfrac{\partial^2 f_2}{\partial x_n \partial x_2}(a) \\ \vdots & \vdots & & \vdots \\ \dfrac{\partial^2 f_n}{\partial x_1 \partial x_n}(a) & \dfrac{\partial^2 f_n}{\partial x_2 \partial x_n}(a) & \cdots & \dfrac{\partial^2 f_n}{\partial x_n^2}(a) \end{bmatrix}$$

que, segundo o Teorema de Schwarz, é uma matriz simétrica.

CAPÍTULO

2

OTIMIZAÇÃO

Otimização é o processo que procura uma solução que forneça o máximo benefício segundo algum critério. Em outras palavras, é a busca de uma condição ótima (a melhor solução).

Problemas de otimização são comuns no nosso dia-a-dia e aparecem, por exemplo, quando procuramos saber qual a quantidade de produção mais econômica para uma empresa, o ponto mais próximo da órbita de um cometa da Terra, as dimensões de um terreno para maximizar sua área, dentre outras situações.

A seguir, trataremos do problema de encontrar máximos e mínimos para funções de várias variáveis reais a valores reais. Este capítulo foi baseado na referência [2].

2.1 Otimização não condicionada

Definição 6. Seja $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ uma função real e seja $(x_0, y_0) \in A$, com $A \subset D$. Dizemos que (x_0, y_0) será ponto de máximo de f em A se, para todo $(x, y) \in A$, $f(x, y) \leq f(x_0, y_0)$. Nesse caso, $f(x_0, y_0)$ é denominado valor máximo de f em A. Dizemos que $(x_0, y_0) \in D$ é ponto de **máximo global ou absoluto de** f se, para todo $(x, y) \in D, f(x, y) \leq f(x_0, y_0)$. Nesse caso, $f(x_0, y_0)$ é denominado valor máximo de f. Dizemos que $(x_0, y_0) \in D$ é ponto de **máximo local** de f se existir uma bola aberta B de centro (x_0, y_0) tal que $f(x, y) \leq f(x_0, y_0)$ para todo $(x, y) \in B \cap D$.

Definição 7. Dizemos que (x_0, y_0) será ponto de mínimo de f em A se, para todo $(x, y) \in A$, $f(x, y) \geq f(x_0, y_0)$. Nesse caso, $f(x_0, y_0)$ é denominado valor mínimo de f em A. Dizemos que $(x_0, y_0) \in D$ é ponto de **mínimo global ou absoluto de** f se, para todo $(x, y) \in D, f(x, y) \geq f(x_0, y_0)$. Nesse caso, $f(x_0, y_0)$ é denominado valor mínimo de f. Dizemos que $(x_0, y_0) \in D$ é ponto de **mínimo local** de f se existir uma bola aberta B de centro (x_0, y_0)

tal que $f(x,y) \geq f(x_0,y_0)$ para todo $(x,y) \in B \cap D$.

Definição 8. Dizemos que $(x_0,y_0) \in D$ é **ponto de sela de** f, se para toda bola aberta B centrada em (x_0,y_0), existem pontos do domínio onde $f(x,y) < f(x_0,y_0)$ e pontos do domínio onde $f(x,y) > f(x_0,y_0)$.

Os pontos de máximo e mínimo de uma função são denominados **extremantes** de f.

Exemplo 27. Mostraremos que a função $f(x,y) = (x-1)^2 + (y-2)^2$ - 1 admite um mínimo global no ponto $P(1,2)$.

Sabemos que $(x-1)^2 \geq 0$ e $(y-2)^2 \geq 0$.

Assim, $(x-1)^2 + (y-2)^2 \geq 0 \Rightarrow (x-1)^2 + (y-2)^2 - 1 \geq -1 \Rightarrow f(x,y) \geq -1 = f(1,2)$, para todo $(x,y) \in \mathbb{R}^2$.

Logo, $(1,2)$ é um ponto de mínimo global.

Exemplo 28. Considere a função $f(x,y) = x^2 - y^2$. Temos que $(0,0)$ é um ponto de sela, pois para qualquer vizinhança de $(0,0)$, podemos encontrar pontos onde $f(x,y) < 0$ e $f(x,y) > 0$, basta considerar $f(x,0) = x^2 > 0$ e $f(0,y) = -y^2 < 0$, para $x \neq 0$ e $y \neq 0$.

Temos os seguintes resultados sobre os extremantes locais,

Teorema 8. Seja (x_0,y_0) um ponto interior de D e vamos supor que $\dfrac{\partial f}{\partial x}(x_0,y_0)$ e $\dfrac{\partial f}{\partial y}(x_0,y_0)$ existam. Sob essas hipóteses, uma condição necessária para que (x_0,y_0) seja um extremante local de f é que $\dfrac{\partial f}{\partial x}(x_0,y_0) = 0$ e $\dfrac{\partial f}{\partial y}(x_0,y_0) = 0$.

Demonstração. Suponhamos que (x_0,y_0) seja um ponto de máximo local. Como $(x_0,y_0) \in D$, existe uma bola aberta $B \subset D$, de centro (x_0,y_0) tal que, para todo (x,y) em B, $f(x,y) \leq f(x_0,y_0)$. Por outro lado, existe um intervalo aberto I, com $x_0 \in I$, tal que para todo $x \in I$, $(x,y_0) \in B$.

Consideremos a função $g : I \subset \mathbb{R} \to \mathbb{R}$, dada por $g(x) = f(x,y_0)$.

Temos:

i) g é derivável em x_0 e $g'(x_0) = \dfrac{\partial f}{\partial x}(x_0,y_0)$;

ii) x_0 pertence ao interior de I;

iii) x_0 é máximo local de g.

Logo, $g'(x_0) = 0$ e, portanto $\dfrac{\partial f}{\partial x}(x_0,y_0) = 0$.

De maneira análoga, consideremos a função $g : J \to \mathbb{R}$, dada por $g(y) = f(x,y)$, onde J é um intervalo aberto contendo y_0 e para todo $y \in J$, $(x_0,y) \in B$.

Temos:

i) g é derivável em y_0 e $g'(y_0) = \dfrac{\partial f}{\partial y}(x_0, y_0)$;

ii) y_0 pertence ao interior de J;

iii) y_0 é máximo local de g.

Logo, $g'(y_0) = 0$ e, portanto $\dfrac{\partial f}{\partial y}(x_0, y_0) = 0$.

A mesma análise é feita para (x_0, y_0) sendo um ponto de mínimo local.

□

Denominamos por **pontos críticos** de uma função f os pontos interiores ao domínio de f em que todas as derivadas parciais de primeira ordem são nulas.

Esse teorema nos indica que se f admite derivadas parciais em todos os pontos interiores do domínio de f, então os únicos candidatos a extremantes locais de f, que são pontos interiores do domínio, são pontos críticos de f.

Exemplo 29. Vamos determinar os pontos críticos da função $f(x,y) = 3xy^2 + x^3 - 3x$.

Primeiramente determinaremos as derivadas parciais:

$$\frac{\partial f}{\partial x} = 3y^2 + 3x^2 - 3 \text{ e } \frac{\partial f}{\partial y} = 6xy.$$

Sabemos que se $\dfrac{\partial f}{\partial x} = 0$ e $\dfrac{\partial f}{\partial y} = 0$, então $\begin{cases} 3y^2 + 3x^2 - 3 = 0 \\ 6xy = 0 \end{cases}$.

Da segunda equação podemos concluir que $x = 0$ ou $y = 0$. Substituindo $x = 0$ na primeira equação, temos $3y^2 - 3 = 0$, o que nos dá $y = \pm 1$.

Assim, temos os pontos críticos $(0, -1)$ e $(0, 1)$.

Substituindo $y = 0$ na primeira equação, temos $3x^2 - 3 = 0$, o que nos dá $x = \pm 1$.

Dessa forma temos os pontos críticos $(-1, 0)$ e $(1, 0)$.

Portanto, a função $f(x,y) = 3xy^2 + x^3 - 3x$ possui quatro pontos críticos, a saber : $(-1,0)$, $(1,0)$, $(0,-1)$ e $(0,1)$.

Observação: Um ponto $(x_0, y_0) \in A$ que não é ponto interior de A é denominado ponto de fronteira de A. O teorema que acabamos de ver não é aplicado a pontos de fronteira do domínio de f, porém esse tipo de ponto pode ser extremante local sem que as derivadas parciais se anulem nele. Os pontos de fronteira devem ser analisados separadamente.

O seguinte Teorema nos auxilia na classificação dos pontos críticos quanto a serem pontos de máximo ou mínimo locais.

Teorema 9. Sejam $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 e (x_0, y_0) um ponto interior de D e $\mathscr{H}f(x_0, y_0)$ a matriz hessiana de f calculada em (x_0, y_0). Suponhamos que (x_0, y_0) seja ponto crítico de f. Então:

- Se $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) < 0$ e $\det(\mathscr{H}f(x_0, y_0)) > 0$, então$(x_0, y_0)$ será ponto de máximo local de f.
- Se $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) > 0$ e $\det(\mathscr{H}f(x_0, y_0)) > 0$, então$(x_0, y_0)$ será ponto de mínimo local de f.
- Se $\det(\mathscr{H}f(x_0, y_0)) < 0$ então (x_0, y_0) será ponto de sela.

Demonstração. Pelo Teorema de Taylor temos que:

$$f(x,y) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + E_1(x,y), \text{ onde}$$

$$E_1(x,y) = \frac{1}{2}\left[\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})(x-x_0)^2 + \frac{2\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y})(x-x_0)(y-y_0) + \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y})(y-y_0)^2\right]$$

Consideremos a forma quadrática $Q(h,k) = ah^2 + 2bhk + ck^2$. Podemos organizá-la da seguinte forma:

$$ah^2 + 2bhk + ck^2 = a\left(h^2 + \frac{2bhk}{a} + \frac{ck^2}{a}\right) = a\left(h^2 + \frac{2b}{a}hk + \frac{b^2k^2}{a^2} - \frac{b^2k^2}{a^2} + \frac{ck^2}{a}\right)$$

$$= a\left[\left(h + \frac{bk}{a}\right)^2 + \frac{ac-b^2}{a^2}k^2\right] = a\left(h + \frac{bk}{a}\right)^2 + \frac{\begin{vmatrix} a & b \\ b & c \end{vmatrix}}{a}k^2.$$

Assim, comparando $2E_1(x,y)$ com $Q(h,k)$, temos:

$$h = (x-x_0),\ k = (y-y_0),\ a = \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}),\ b = \frac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) \text{ e } c = \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y}).$$

Então podemos escrever:

$$2E_1(x,y) =$$

$$\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})\left[(x-x_0) + \frac{\frac{\partial^2 f}{\partial x + \partial y}(\bar{x},\bar{y})(y-y_0)}{\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})}\right]^2 + \frac{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}) & \frac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) \\ \frac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y}) \end{vmatrix}}{\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y})}(y-y_0)^2.$$

Pelo Teorema da Conservação do Sinal, existe uma bola aberta $B \subset D$ (pois (x_0, y_0) é ponto interior de D, as derivadas de segunda ordem de f são contínuas, $f_{xx}(x_0, y_0) < 0$ e

$\det(\mathscr{H}f(x_0,y_0)) > 0)$ e de centro (x_0,y_0), tal que, para todo (x,y) em B, $\dfrac{\partial^2 f}{\partial x^2}(x,y) < 0$ e $\det(\mathscr{H}f(x,y)) > 0$. Podemos supor que $(\bar{x},\bar{y}) \in B$.

Como já sabemos, as derivadas parciais de primeira ordem são nulas, $\dfrac{\partial f}{\partial x}(x_0,y_0) = \dfrac{\partial f}{\partial y}(x_0,y_0) = 0$, como $\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}) < 0$ e $\begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}) & \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) \\ \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) & \dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y}) \end{vmatrix} > 0$, então, $E_1(x,y) < 0$, logo $f(x,y) < f(x_0,y_0)$, para $(x,y) \in B$. Portanto $f(x_0,y_0)$ é ponto de máximo local.

Da mesma forma, existe $B_1 \subset D$ tal que, para todo (x,y) em B_1, $\dfrac{\partial^2 f}{\partial x^2}(x,y) > 0$ e $\det(\mathscr{H}f(x,y)) > 0$.

Podemos supor que $(\bar{x},\bar{y}) \in B_1$ e temos $\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}) > 0$ e $\begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y}) & \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) \\ \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y}) & \dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y}) \end{vmatrix} > 0$, então, $E_1(x,y) > 0$, logo $f(x,y) > f(x_0,y_0)$, para $(x,y) \in B_1$. Portanto, $f(x_0,y_0)$ é ponto de mínimo local.

Para mostrar o último item, considere $g(t) = f(x_0 + ht, y_0 + kt)$, $t \in [0,1]$ isto é, f restrita ao segmento de reta na direção do vetor $\vec{v} = (h,k)$.

Pela regra da cadeia

$$g'(t) = \frac{\partial f}{\partial x}(x_0+ht,y_0+kt)h + \frac{\partial f}{\partial y}(x_0+ht,y_0+kt)k$$

e

$$g''(t) = \left[\frac{\partial^2 f}{\partial x^2}(x_0+ht,y_0+kt)h + \frac{\partial^2 f}{\partial y \partial x}(x_0+ht,y_0+kt)k\right] h \; +$$
$$+ \left[\frac{\partial^2 f}{\partial x \partial y}(x_0+ht,y_0+kt)h + \frac{\partial^2 f}{\partial y^2}(x_0+ht,y_0+kt)k\right] k$$

e assim

$$g''(0) = \frac{\partial^2 f}{\partial x^2}(x_0,y_0)h^2 + \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)hk + 2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)k^2.$$

Como $\det(\mathscr{H}f(x_0,y_0)) < 0$, considerando $(h_1,k_1) = (1,0)$ e $(h_2,k_2) = \left(\dfrac{\dfrac{\partial^2 f}{\partial x \partial y}(x_0,y_0)}{\dfrac{\partial^2 f}{\partial x^2}(x_0,y_0)}, -1\right)$, teremos

$$g''(0) = \frac{\partial^2 f}{\partial x^2}(x_0,y_0) \text{ e } g''(0) = \frac{det(\mathscr{H}f(x_0,y_0))}{\dfrac{\partial^2 f}{\partial x^2}(x_0,y_0)}$$

que terão sinais contrários, logo, (x_0,y_0) será ponto de máximo (mínimo) na direção de (h_1,k_1) e ponto de mínimo (máximo) na direção de (h_2,k_2), dependendo do sinal de

$\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0)$.

Se $\dfrac{\partial^2 f}{\partial x^2}(x_0, y_0) = 0$, como $\det(\mathscr{H} f(x_0, y_0)) < 0$, temos $\dfrac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \neq 0$. Podemos encontrar α real tal que $(h_1, k_1) = (\alpha, 1)$ e $(h_2, k_2) = (\alpha, -1)$ fornecem sinais contrários.

Portanto $f(x_0, y_0)$ é ponto de sela.

$\square$

Faremos o mesmo raciocínio para uma função de três variáveis.

Sendo $2E_1(x,y,z) = \dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})(x-x_0)^2 + \dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z})(y-y_0)^2 + \dfrac{\partial^2 f}{\partial z^2}(\bar{x},\bar{y},\bar{z})(z-z_0)^2 +$
$+2\dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y},\bar{z})(x-x_0)(y-y_0) + 2\dfrac{\partial^2 f}{\partial x \partial z}(\bar{x},\bar{y},\bar{z})(x-x_0)(z-z_0) + 2\dfrac{\partial^2 f}{\partial y \partial z}(\bar{x},\bar{y},\bar{z})(y-y_0)(z-z_0)$, e considerando a forma quadrática

$Q(r,s,t) = ar^2 + bs^2 + ct^2 + 2drs + 2ert + 2fst$, podemos organizá-la da seguinte forma:

$$ar^2 + bs^2 + ct^2 + 2drs + 2ert + 2fst = a\left(r^2 + \frac{b}{a}s^2 + \frac{c}{a}t^2 + \frac{2d}{a}rs + \frac{2e}{a}rt + \frac{2f}{a}st\right) =$$

$$a\left(r^2 + \frac{d^2}{a^2}s^2 + \frac{e^2}{a^2}t^2 + \frac{2d}{a}rs + \frac{2e}{a}rt + \frac{2e}{a^2}dst - \frac{d^2}{a^2}s^2 - \frac{e^2}{a^2}t^2 - \frac{2e}{a^2}dst + \frac{2f}{a}st + \frac{b}{a}s^2 + \frac{c}{a}t^2\right) =$$

$$a\left[\left(r + \frac{d}{a}s + \frac{e}{a}t\right)^2 + \left(\frac{ab-d^2}{a^2}\right)s^2 + \left(\frac{ac-e^2}{a^2}\right)t^2 + \frac{2(af-ed)}{a^2}st\right] =$$

$$a\left(r + \frac{d}{a}s + \frac{e}{a}t\right)^2 + \frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}s^2 + \frac{\begin{vmatrix} a & e \\ e & c \end{vmatrix}}{a}t^2 + 2\frac{\begin{vmatrix} a & e \\ d & f \end{vmatrix}}{a}st$$

Assim,

$$Q(r,s,t) = a\left(r + \frac{d}{a}s + \frac{e}{a}t\right)^2 + \frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}\left(s^2 + \frac{\begin{vmatrix} a & e \\ e & c \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}t^2 + 2\frac{\begin{vmatrix} a & e \\ d & f \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}st\right) =$$

$$a\left(r+\frac{d}{a}s+\frac{e}{a}t\right)^2+\frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}\left(s^2+2\frac{\begin{vmatrix} a & e \\ d & f \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}st+\frac{\begin{vmatrix} a & e \\ e & f \end{vmatrix}^2}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}^2}t^2-\frac{\begin{vmatrix} a & e \\ e & f \end{vmatrix}^2}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}^2}t^2+\frac{\begin{vmatrix} a & e \\ e & c \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}t^2\right)=$$

$$a\left(r+\frac{d}{a}s+\frac{e}{a}t\right)^2+\frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}\left[\left(s+\frac{\begin{vmatrix} a & e \\ e & f \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}t\right)^2+\frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}\begin{vmatrix} a & e \\ e & c \end{vmatrix}-\begin{vmatrix} a & e \\ e & f \end{vmatrix}^2 t^2}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}\right]=$$

$$a\left(r+\frac{d}{a}s+\frac{e}{a}t\right)^2+\frac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}\left[\left(s+\frac{\begin{vmatrix} a & e \\ e & f \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}t\right)^2+at^2\frac{\begin{vmatrix} a & d & e \\ d & b & f \\ e & f & c \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}^2}\right]$$

Portanto, $Q(r,s,t)=a\left(r+\dfrac{ds}{a}+\dfrac{et}{a}\right)^2+\dfrac{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}{a}\left[\left(s+\dfrac{\begin{vmatrix} a & e \\ e & f \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}t\right)^2+at^2\dfrac{\begin{vmatrix} a & d & e \\ d & b & f \\ e & f & c \end{vmatrix}}{\begin{vmatrix} a & d \\ d & b \end{vmatrix}}\right].$

Assim, comparando $2E_1(x,y,z)$ com $Q(r,s,t)$, temos:

$r=(x-x_0)$, $s=(y-y_0)$, $t=(z-z_0)$, $a=\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})$, $b=\dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z})$, $c=\dfrac{\partial^2 f}{\partial z^2}(\bar{x},\bar{y},\bar{z})$, $d=\dfrac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z})$, $e=\dfrac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z})$ e $f=\dfrac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z})$.

Então, podemos escrever:

$$2E_1(x,y,z)=\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})\left((x-x_0)+\frac{\dfrac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z})}{\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})}(y-y_0)+\frac{\dfrac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z})}{\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})}(z-z_0)\right)^2$$

$$+\frac{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix}}{\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z})}\left((y-y_0)\frac{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z}) \end{vmatrix}}{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix}}(z-z_0)\right)^2 +$$

$$+\frac{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial z^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix}}{\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix}}(z-z_0)^2.$$

Usando os mesmos argumentos para o caso $n=2$, se $\frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) \leq 0$,

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix} \geq 0$$

e

$$\begin{vmatrix} \frac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial y}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z}) \\ \frac{\partial^2 f}{\partial x\partial z}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial y\partial z}(\bar{x},\bar{y},\bar{z}) & \frac{\partial^2 f}{\partial z^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix} \leq 0,$$

e, portanto, $E_1(x,y,z) \leq 0$. Logo, $f(x,y,z) = f(x_0,y_0,z_0) + \frac{1}{2}E_1(x,y,z)$, o que implica que

$$f(x,y,z) - f(x_0,y_0,z_0) = \frac{1}{2}E_1(x,y,z) \leq 0$$

e portanto,

$$f(x,y,z) \leq f(x_0,y_0,z_0),\text{para } (x,y,z) \in B \subset D,$$

ou seja, $f(x_0,y_0,z_0)$ é ponto de máximo local .

Por outro lado, novamente, usando os argumentos citados para o caso $n=2$, se $\dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) \geq 0$,

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y},\bar{z}) \\ \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix} \geq 0$$

e

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial x^2}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial x \partial z}(\bar{x},\bar{y},\bar{z}) \\ \dfrac{\partial^2 f}{\partial x \partial y}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial y^2}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial y \partial z}(\bar{x},\bar{y},\bar{z}) \\ \dfrac{\partial^2 f}{\partial x \partial z}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial y \partial z}(\bar{x},\bar{y},\bar{z}) & \dfrac{\partial^2 f}{\partial z^2}(\bar{x},\bar{y},\bar{z}) \end{vmatrix} \geq 0,$$

então $E_1(x,y,z) \geq 0$. Logo,

$$f(x,y,z) - f(x_0,y_0,z_0) \geq 0 \Rightarrow f(x,y,z) \geq f(x_0,y_0,z_0),$$

para $(x,y,z) \in B_1 \subset D$. Portanto, $f(x_0,y_0,z_0)$ é ponto de mínimo local.

Para os exemplos anteriores temos:

1) Se $f(x,y) = (x-1)^2 + (y-2)^2 - 1$, então o único ponto crítico é $(1,2)$.

Temos que $f_{xx} = (1,2) = 2 > 0$ e $\det(\mathscr{H} f(1,2)) = 4.1.2 = 8 > 0$. Logo, $(1,2)$ é um ponto de mínimo local.

2) Se $f(x,y) = x^2 - y^2$, então o único ponto crítico é $(0,0)$ e $\det(\mathscr{H} f(0,0)) = -4 < 0$ e, pelo teorema, $(0,0)$ é um ponto de sela.

Podemos perceber que esses resultados estão associados a submatrizes da matriz Hessiana da função f. A matriz Hessiana é um exemplo de uma matriz quadrática simétrica e o estudo destas matrizes auxilia não só na análise de máximos e mínimos locais como também os globais, como veremos mais adiante. Por isso, faremos um estudo sobre tais matrizes.

2.1.1 Representação Matricial de Formas Quadráticas e Classificação de uma Matriz através dos Menores Principais Líderes

[2] As formas quadráticas, assim como as lineares, têm representação matricial, de modo que estudar as propriedades de uma forma quadrática é o mesmo que estudar as propriedades de uma matriz simétrica. Podemos escrever a forma quadrática $a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2$ como matriz de diferentes formas, dependendo de como repartimos o coeficiente do termo misto. A maneira ideal é dividí-lo igualmente, de modo a obter uma matriz simétrica:

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{22}x_2^2 =$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a_{11} & 1/2a_{12} \\ 1/2a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Para a forma quadrática em três variáveis, podemos escrever:

$$a_{11}x_1^2 + a_{12}x_1x_2 + a_{13}x_1x_3 + a_{22}x_2^2 + a_{23}x_2x_3 + a_{33}x_3^2 =$$

$$\begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{bmatrix} a_{11} & 1/2a_{12} & 1/2a_{13} \\ 1/2a_{12} & a_{22} & 1/2a_{23} \\ 1/2a_{13} & 1/2a_{23} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Por indução finita, temos que a forma quadrática $\sum\limits_{i \leq j} a_{ij}x_ix_j$ pode ser escrita como:

$$\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & \frac{1}{2}a_{12} & \dots & \frac{1}{2}a_{1n} \\ \frac{1}{2}a_{21} & a_{22} & \dots & \frac{1}{2}a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2}a_{n1} & \frac{1}{2}a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix},$$

ou ainda, x^TAx, onde A é uma matriz simétrica.

Uma matriz simétrica é considerada positiva, negativa, não negativa, não positiva ou indefinida, se a forma quadrática $Q(x) = x^TAx$ for positiva, negativa, não negativa, não positiva ou indefinida, respectivamente, segundo a seguinte definição,

Definição 9. Seja A uma matriz simétrica de ordem n, então A é:

i) Positiva se $x^TAx > 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$;

ii) Não negativa se $x^TAx \geq 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$;

iii) Negativa se $x^TAx < 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$;

iv) Não positiva se $x^TAx \leq 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$;

v) Indefinida se $x^TAx > 0$, para alguns $x \in \mathbb{R}^n$ e $x^TAx < 0$ para outros $x \in \mathbb{R}^n$.

As matrizes e formas quadráticas positivas e negativas são denominadas definidas, enquanto as não negativas e não positivas são denominadas semidefinidas.

Exemplo 30. Seja $f(x,y) = 2x^2 + 2y^2$ a forma quadrática associada à matriz simétrica é $A = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$. Obviamente, $f(x,y) > 0$, $\forall (x,y) \neq (0,0)$. Logo, A é positiva.

Exemplo 31. Seja, em $\mathbb{R}^3$, a forma quadrática $f(x,y,z) = 3x^2 + 2y^2 - 5z^2 + 4xy + 2xz - yz$, associada à matriz simétrica $A = \begin{bmatrix} 3 & 2 & 1 \\ 2 & 2 & -\frac{1}{2} \\ 1 & -\frac{1}{2} & -5 \end{bmatrix}$.

A soma das parcelas que contêm x se escreve como

$$\begin{aligned} 3x^2 + 4xy + 2xz &= 3\left[x^2 + 2x\left(\frac{2}{3}y + \frac{1}{3}z\right)\right] \\ &= 3\left[\left(x + \frac{2}{3}y + \frac{1}{3}z\right)^2 - \left(\frac{2}{3}y + \frac{1}{3}z\right)^2\right] \\ &= 3\left[u^2 - \frac{1}{3}(2y+z)^2\right], \text{ onde } u = x + \frac{2}{3}y + \frac{1}{3}z. \end{aligned}$$

Portanto $f(x,y,z) = 3u^2 - \frac{1}{3}[(2y+z)]^2 + 2y^2 - 5z^2 - yz = 3u^2 + \frac{2}{3}y^2 - \frac{16}{3}z^2 - \frac{7}{3}yz$.

A soma das parcelas que contêm y nessa última expressão é

$\frac{2}{3}y^2 - \frac{7}{3}yz = \frac{2}{3}\left(y^2 - 2y\frac{7}{4}z\right) = \frac{2}{3}\left[\left(y - \frac{7}{4}z\right)^2 - \frac{49}{16}z^2\right] = \frac{2}{3}v^2 - \frac{49}{16}z^2,$

onde $v = y - \frac{7}{4}z$.

Portanto,

$$f(x,y,z) = 3u^2 + \frac{2}{3}v^2 - \frac{49}{24}z^2 - \frac{16}{3}z^2 = 3u^2 + \frac{2}{3}v^2 - \frac{177}{24}z^2.$$

Isso mostra que f é indefinida, pois assume valores positivos quando $z = 0$ e $u^2 + v^2 \neq 0$ e valores negativos, quando $u = v = 0$ e $z \neq 0$.

A classificação de uma matriz simétrica é de grande importância na teoria da otimização, bem como a concavidade de uma função, pois o teste de segunda derivada em $\mathbb{R}^n$ envolve analisar se a matriz da derivada segunda, a matriz hessiana, de uma função em um ponto crítico é positiva, negativa ou indefinida.

Veremos agora, como classificar matrizes como positiva, negativa etc, através de seus menores principais.

Definição 10. Seja A uma matriz $n \times n$. Uma submatriz principal de ordem k de A é uma submatriz de A de tamanho $k \times k$ formada a partir de A suprimindo n - k colunas, digamos, as colunas i_1, i_2, ..., i_{n-k} e as mesmas n - k linhas, ou seja, as linhas i_1, i_2, ..., i_{n-k}.

O determinante de uma submatriz principal $k \times k$ é denominado de menor principal de ordem k de A.

Por exemplo, para uma matriz de ordem 3 qualquer

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix},$$

existe um menor principal de 3^a ordem, que é $detA$, existem 3 menores principais de 2^a ordem

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \text{ e } \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix},$$

e existem 3 menores principais de 1^a ordem

$$\begin{vmatrix} a_{11} \end{vmatrix}, \begin{vmatrix} a_{22} \end{vmatrix} \text{ e } \begin{vmatrix} a_{33} \end{vmatrix}.$$

Observação: Nenhuma outra submatriz de A é uma submatriz principal.

Dentre os menores principais de ordem k de uma matriz dada, existe um em especial que devemos destacar.

Definição 11. Seja A uma matriz n x n. A submatriz principal de ordem k de A obtida suprimindo as últimas n - k linhas e as últimas n - k colunas de A é denominada submatriz principal líder de ordem k em A. Seu determinante é denominado menor principal líder de k em A. Vamos denotar a submatriz principal líder de ordem k por A_k e o correspondente menor principal líder por $|A_k|$.

Uma matriz $n \times n$ tem n submatrizes principais líderes: a submatriz de ordem 1, mais a esquerda e acima, a submatriz de ordem 2, mais a esquerda e acima e assim sucessivamente. Para a matriz de ordem 3 geral, analisada acima, os 3 menores principais líderes são:

$$\begin{vmatrix} a_{11} \end{vmatrix}, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \text{ e } \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}.$$

Enunciaremos dois lemas referentes à matrizes simétricas (ver [2],p.402), que serão utilizados na próxima demonstração, cujas demonstrações podem ser encontradas em [2], p.402 e 403.

Lema 2. Se A é uma matriz positiva ou negativa, então A é não singular.

Lema 3. Suponha que A é uma matriz simétrica e que Q é uma matriz não singular. Então, Q^TAQ é uma matriz simétrica e A é positiva/negativa se, e somente se, Q^TAQ é positiva/negativa.

Teorema 10. Seja A uma matriz quadrada de ordem n simétrica. Então,

(a) A é positiva se, e somente se, todos os n menores principais líderes de A são (estritamente) positivos;

(b) A é negativa se, e somente se, os n menores principais líderes de A alternam o sinal, da seguinte forma: $|A_1| < 0$, $|A_2| > 0$, $|A_3| < 0$, ... O k-ésimo menor principal líder deve ter o mesmo sinal de $(-1)^k$.

(c) Se algum menor principal líder de A de ordem k (ou um par de menores) é não nulo mas não se encaixa em nenhum dos dois padrões de sinal acima, então A é indefinida. Esse caso ocorre quando A tem um menor principal líder de ordem k negativo com um k inteiro par ou quando A tem um menor principal líder positivo de ordem l, com k e l dois inteiros ímpares distintos.

Demonstração. Faremos apenas a demonstração do Item (a)

Para matrizes de ordem 1, há apenas um menor principal líder que é o próprio determinante da matriz e como o determinante de uma matriz unitária é o próprio elemento, se este for positivo, a matriz é positiva, e vice-versa.

Para matrizes de ordem 2, são 2 menores principais líderes, o de ordem 1 e o de ordem 2.

Seja $A = \begin{bmatrix} a & \frac{b}{2} \\ \frac{b}{2} & c \end{bmatrix}$, então o menor principal líder de ordem 1 é $|a| = a$ e o menor principal líder de ordem 2 é o determinante da própria matriz, dado por $ac - \frac{b^2}{4} = \frac{4ac-b^2}{4}$.

Como vimos anteriormente, por definição uma matriz simétrica é positiva se $x^TAx > 0$, $\forall x \neq 0$, $x \in \mathbb{R}^n$.

Assim, se A é positiva, $x^TAx > 0$, ou seja

$$ax^2 + \frac{b}{2}xy + \frac{b}{2}xy + cy^2 > 0, \text{ isto é,}$$

$$ax^2 + bxy + cy^2 > 0, \ \forall (x,y) \neq (0,0).$$

Fixando $y = y_0$ qualquer e sendo x a variável da equação $ax^2 + bxy_0 + cy_0^2 = 0$, temos:

$$x = \frac{-by_0 \pm |y_0|\sqrt{b^2 - 4.a.c}}{2a}.$$

Se $a > 0$ e $4ac - b^2 > 0$, temos $ax^2 + 2bxy_0 + cy_0^2 > 0$, para todo y_0.

Logo, se os menores principais líderes são positivos, então a matriz é positiva.

Suponha que A é positiva, logo $ax^2 + bxy + cy^2 > 0$, $\forall (x,y) \neq (0,0)$.

Se $a > 0$ e $4ac - b^2 < 0$, então existe y_0 tal que $ax^2 + bxy_0 + cy_0^2 < 0$.

Se $a < 0$ e $4ac - b^2 < 0$, então existe y_0 tal que $ax^2 + bxy_0 + cy_0^2 < 0$ (por exemplo (x_0, y_0) com x_0 maior do que a raiz).

Se $a < 0$ e $4ac - b^2 > 0$, então $ax^2 + bxy_0 + cy_0^2 < 0$, para todo y_0.

Logo, se A é positiva é necessário que $a > 0$ e $4ac - b^2 > 0$.

Desta forma, A é positiva se, e somente se, os menores principais líderes são positivos.

Analizaremos agora, matrizes de ordem 3.

Por indução vale para $n = 2$, se A_2 é positiva, então $|A_1|$ e $|A_2|$ são positivos.

Seja

$$A_3 = \begin{pmatrix} a & b & a_{13} \\ b & c & a_{23} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}.$$

Podemos reescrevê-la da seguinte forma:

$$A_3 = \underbrace{\begin{pmatrix} I_2 & 0 \\ \left(A_2^{-1}\begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix}\right)^T & 0 \\ & 1 \end{pmatrix}}_{Q^T} \underbrace{\begin{pmatrix} A_2 & 0 \\ & 0 \\ 0 \quad 0 & d \end{pmatrix}}_{B} \underbrace{\begin{pmatrix} I_2 & A_2^{-1}\begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} \\ & \\ 0 \quad 0 & 1 \end{pmatrix}}_{Q}$$

onde $I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $A_2 = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ e $d = a_{3,3} - a^T A_2^{-1} a$

Por propriedade dos determinantes, $detQ = detQ^T = 1$ e $detB = ddetA_2$.

Estamos supondo que todos os menores principais líderes são positivos, então $detA_2 > 0$ e $detA_3 = detA > 0$ e , então $d > 0$. Mostraremos que B é positiva. De fato, se

$$\begin{pmatrix} x^T & x_3 \end{pmatrix} \begin{pmatrix} A_2 & 0 \\ & 0 \\ 0 \quad 0 & d \end{pmatrix} \begin{pmatrix} x \\ x_3 \end{pmatrix} =$$

$$= x^T A_2 x + dx_3^2,$$

então, como $x^T A_2 x > 0$, pela hipótese de indução e $x_3^2 > 0$, segue que B é positiva. Pelo Lema 3, A_3 é positiva.

Logo, se os menores principais líderes são positivos, então a matriz é positiva.

Agora, vamos supor A positiva.

Assim, B é positiva e $x^T A_2 x + dx_3^2 > 0$. Como A_2 é positiva, por indução, temos $x^T A_2 x > 0$ e então $dx_3^2 > 0$, dessa forma $d > 0 \Rightarrow detA_3 =$ d.$detA_2 > 0$.

Logo, se a matriz é positiva, então os menores principais líderes são positivos.

Agora, supondo que o teorema é verdadeiro para matrizes de ordem n vamos prová-lo para matrizes de ordem $n+1$.

Sejam uma matriz de ordem $n+1$ simétrica e A_j a submatriz principal líder de A, para $j = 1,...,n+1$.

Primeiramente, se todos os A_j têm determinante positivo, então A é positiva. As submatrizes principais líderes de A_n são $A_1,...,A_n$, que são positivas por hipótese de indução, já que são as primeiras n submatrizes principais líderes de A. Pela hipótese de indução, a saber, que o teorema é verdadeiro para matrizes de ordem n, a matriz n x n simétrica A_n é positiva. Como já visto, A_n é invertível.

Escrevendo $A = \begin{pmatrix} A_n & a \\ a^T & a_{n+1,n+1} \end{pmatrix}$, onde $a = \begin{pmatrix} a_{1,n+1} \\ \vdots \\ a_{n,n+1} \end{pmatrix}$,

Considerando $d = a_{n+1,n+1} - a^T A_n^{-1} a$, I_n a matriz identidade de ordem n e 0_n a matriz coluna n x 1 nula. Então, a matriz A pode ser escrita como:

$$A = \begin{pmatrix} I_n & 0_n \\ (A_n^{-1}a)^T & 1 \end{pmatrix} \begin{pmatrix} A_n & 0_n \\ 0_n^T & d \end{pmatrix} \begin{pmatrix} I_n & A_n^{-1}a \\ 0_n^T & 1 \end{pmatrix},$$

$$A = Q^T B Q.$$

Por propriedade dos determinantes, $detQ = detQ^T = 1$ e $detB = ddetA_n$.

Portanto , $detA =$d$detA_n$. Como $detA > 0$ e $detA_n > 0$, temos $d > 0$.

Seja X um vetor $(n+1)$-dimensional arbitrário. Escrevendo $X = \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}$, onde x é um vetor n-dimensional. Então,

$X^T BX = \begin{pmatrix} x^T & x_{n+1} \end{pmatrix} \begin{pmatrix} A_n & 0_n \\ 0_n & d \end{pmatrix} \begin{pmatrix} x \\ x_{n+1} \end{pmatrix}$ e

$X^T BX = x^T A_n x + dx_{n+1}^2$

Como A_n é positiva, e $d > 0$, $x^TAx + dx_{n+1}^2$ é estritamente positiva. Portanto, $B = Q^TAQ$ é positiva. Dessa forma, A é positiva.

Para provar a recíproca, ou seja, que A positiva implica que todos os $|A_j|$ são positivos, novamente usamos indução. Vimos que este resultado é verdadeiro para matrizes de ordem 1 e ordem 2. Considere que também é verdadeiro para matrizes simétricas de ordem n e seja A uma matriz de ordem $(n+1)$ simétrica positiva. Primeiro mostraremos que todas as A_j são positivas. Seja x_j um vetor não nulo de dimensão j e seja O^* o vetor nulo de dimensão $(n+1)-j$. Então,

$$0 < \begin{pmatrix} x_j^T & 0^* \end{pmatrix} A \begin{pmatrix} x_j \\ 0^* \end{pmatrix} = x_j^T A_j x_j$$

e A_j é positiva.

Em particular, como A_n é positiva, a hipótese de indução garante que todos $A_1, \ldots, A_n$ têm determinante positivo. Resta provar que o determinante da própria A é positvo. Como A_n é invertível, podemos novamente escrever $A = Q^TBQ$, como anteriormente, e concluir que $d > 0$. Como A é positiva, B é positiva pelo Lema 3. Escolhendo X de tal modo que $x = \vec{0}$ e $x_{n+1} = \vec{1}$. Então,

$$0 < X^TBX = d.$$

Como $detA_n > 0$ e $d > 0$, temos $detA > 0$ □

Voltando aos exemplos anteriores:

1) $f(x,y) = 2x^2 + 2y^2$ é associada à matriz $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, onde os menores principais são $|2| = 2$ e $\begin{vmatrix} 2 & 0 \\ 0 & 2 \end{vmatrix} = 4$, que são positivos. Logo, A é positiva.

2) $f(x,y,z) = 3x^2 + 2y^2 - 5z^2 + 4xz + 2xz - yz$ é associada à matriz $A = \begin{pmatrix} 3 & 2 & 1 \\ 2 & 2 & -1/2 \\ 1 & -1/2 & -5 \end{pmatrix}$, onde os menores principais líderes são $|3| = 3$, $\begin{vmatrix} 3 & 2 \\ 2 & 2 \end{vmatrix} = 2$ e $\begin{vmatrix} 3 & 2 & 1 \\ 2 & 2 & -1/2 \\ 1 & -1/2 & -5 \end{vmatrix} = -\frac{59}{4} < 0$, e pelo Teorema, A é indefinida.

O teste dos menores principais líderes do teorema anterior pode falhar, por exemplo, quando uma dada matriz simétrica A tem algum menor principal líder igual a zero, mas todos os não nulos se encaixam no padrão de sinais dos itens (a) ou (b) do teorema. Quando isso ocorre, a matriz A não é definida, podendo ser semidefinida ou não. Neste caso, para conferir se a matriz é semidefinida, não contamos mais com a possibilidade de conferir apenas o sinal de cada menor principal de A, usando o teste descrito no seguinte teorema. ([2],p.392)

Teorema 11. Seja A uma matriz n x n simétrica. Então, A é não negativa se, e somente se, todos os menores principais líderes de A são maiores ou iguais a zero; A é não positiva se, e somente se, cada menor principal de A de ordem ímpar é menor ou igual a zero e cada menor princpal líder de A de ordem par é maior ou igual a zero.

Podemos, agora, reescrever o Teorema 9 da seguinte forma:

Sejam $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 e (x_0, y_0) um ponto interior de D e $\mathscr{H} f(x_0, y_0)$ a matriz hessiana de f calculada em (x_0, y_0). Suponhamos que (x_0, y_0) seja ponto crítico de f. Então:

- Se a matriz hessiana de f em (x_0, y_0) for positiva, então(x_0, y_0) será ponto de mínimo local de f.
- Se a matriz hessiana de f em (x_0, y_0) for negativa, então(x_0, y_0) será ponto de máximo local de f.
- Se a matriz hessiana de f em (x_0, y_0) for indefinida, então(x_0, y_0) será ponto de sela f.

Generalizando para uma função de n variáveis $f(x_1, x_2, \ldots, x_n)$ teremos que:

- $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ será ponto de máximo local, se $\dfrac{\partial^2 f}{\partial x_1^2}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n) \leq 0$ e todos os demais menores principais da matriz Hessiana de ordem n alternarem o sinal a partir de $\dfrac{\partial^2 f}{\partial x_1^2}(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$.
- $(\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)$ será ponto de mínimo local se todos os demais menores principais da matriz Hessiana de ordem n forem maiores que zero.

2.1.2 Funções Côncavas e Convexas

Definição 12. Seja $f : A \subset \mathbb{R}^n \to \mathbb{R}$, onde A é um subconjunto convexo de $\mathbb{R}^n$, f é côncava se, para quaisquer x, y em A e para todo $t \in [0, 1]$,

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y).$$

Uma função é dita convexa se satisfaz a seguinte condição:

$$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y), \forall t \in [0,1].$$

Em outras palavras, uma função de n variáveis é côncava se, e somente se, qualquer segmento de reta secante ligando dois pontos do gráfico de f fica abaixo dele. Uma função é convexa se, e somente se, qualquer segmento de reta secante ligando dois pontos no seu gráfico fica acima dele.

Teorema 12. Seja $f : A \subset \mathbb{R}^n \to \mathbb{R}$, onde A é um subconjunto convexo de $\mathbb{R}^n$. Então, f é uma função côncava (convexa) se, e somente se, sua restrição a qualquer segmento de reta em A é uma função côncava (convexa) de uma variável.

Demonstração. Suponha que a retrição de f a qualquer segmento de reta em A é uma função côncava. Para provar que f é uma função côncava em A, sejam x e y dois pontos arbitrários de A. Seja $g(t) = f(tx+(1-t)y)$. Por hipótese, g é côncava. Assim, para t entre 0 e 1 temos,

$$
\begin{aligned}
f(tx+(1-t)y) &= g(t) && \text{(definição de } g\text{)}\\
f(tx+(1-t)y) &= g(t.1+(1-t).0) &&\\
f(tx+(1-t)y) &\geq tg(1)+(1-t)g(0) && \text{(pois g é côncava)}\\
f(tx+(1-t)y) &= tf(x)+(1-t)f(y) && \text{(definição de } g\text{)}
\end{aligned}
$$

Consequentemente, f é côncava.

Reciprocamente, suponha que f é côncava. Queremos mostrar que é côncava a restrição $g(t) = f(tx+(1-t)y)$ de f ao segmento de reta contendo x e y. Para fazer isso, fixe s_1 e s_2 e tome t entre 0 e 1. Então,

$$
\begin{aligned}
g(ts_1+(1-t)s_2) &= f((ts_1+(1-t)s_2)x+(1-(ts_1+(1-t)s_2))y) \quad \text{(definição de } g\text{)}\\
g(ts_1+(1-t)s_2) &= f(t(s_1x+(1-s_1)y) + (1-t)(s_2x+(1-s_2)y))\\
g(ts_1+(1-t)s_2) &\geq tf(s_1x+(1-s_1)y) + (1-t)f(s_2x+(1-s_2)y) \quad \text{(concavidade de f)}\\
g(ts_1+(1-t)s_2) &\geq tg(s_1) + (1-t)g(s_2).
\end{aligned}
$$

Portanto, g é côncava.

A prova para funções convexas é análoga.

□

Vamos agora descrever as propriedades desejáveis das funções côncavas e convexas.

2.1.3 Critério para determinação da concavidade de uma função

Ao estudar função de uma variável, temos dois testes analíticos simples para a concavidade.

1) Uma função C^1 em um intervalo I é côncava se, e somente se, sua derivada primeira $f'(x)$ é uma função decrescente de x , para x em I;

2) Uma função C^2 é côncava em um intervalo I se, e somente se, sua derivada segunda $f''(x)$ é menor ou igual a zero para x em I.

Veremos agora um teorema que fornece um resultado bastante relacionado para a concavidade em $\mathbb{R}$ que tem uma generalização óbvia para funções em $\mathbb{R}^n$.

Teorema 13. Seja f uma função de classe C^1 em um intervalo I de $\mathbb{R}$. Então f é côncava em I se, e somente se,

$$f(y)-f(x) \leq f'(x).(y-x), \forall x,y \in \mathbb{R}. \tag{2.1}$$

A função f é convexa em I se, e somente se,

$$f(y)-f(x) \geq f'(x).(y-x), \forall x,y \in \mathbb{R}.$$

Observação 2. Vejamos que a condição (2.1) significa que $f'(x)$ é uma função decrescente. De fato,

$$f(y)-f(x) \leq f'(x).(y-x)$$

$$\frac{f(y)-f(x)}{y-x} \leq f'(x), \tag{2.2}$$

para $x,y \in I$ tais que $y > x$

e

$$\frac{f(y)-f(x)}{y-x} \geq f'(x), \tag{2.3}$$

para $x,y \in I$ tais que $y < x$.

Para vermos que (2.2) e (2.3) implicam que f' é decrescente, suponha $z_1 < z_2$ em I. Então,

$$f'(z_1) \geq \frac{f(z_2)-f(z_1)}{z_2-z_1} = \frac{f(z_1)-f(z_2)}{z_1-z_2}, \text{ para } x = z_1 \text{ e } y = z_2, \text{ em (2.2)}$$

$$\frac{f(z_1)-f(z_2)}{z_1-z_2} = f'(z_2), \text{ com } x = z_2 \text{ e } y = z_1 \text{ em (2.3)}$$

Logo,

$$f'(z_1) \geq f'(z_2).$$

Demonstração. Suponha que uma função f é côncava em I. Sejam $x, y \in I$ e seja $t \in (0,1)$. Então,

$$tf(y) + (1-t)f(x) \leq f(ty + (1-t)x)$$

ou

$$f(y) - f(x) \leq \frac{f(x + t(y-x)) - f(x)}{t}$$

$$f(y) - f(x) \leq \frac{f(x + t(y-x)) - f(x)}{t(y-x)}(y-x).$$

A condição (2.1) segue fazendo $t \to 0$ na última expressão.

Por outro lado, suponha que (2.1) vale para quaisquer x e $y \in I$. Então,

$$f(x) - f((1-t)x + ty) \leq f'((1-t)x + ty)(x - ((1-t)x + ty))$$

$$f(x) - f((1-t)x + ty) \leq f'((1-t)x + ty)(x - x + tx - ty)$$

$$f(x) - f((1-t)x + ty) \leq -tf'((1-t)x + ty)(y-x).$$

Da mesma forma,

$$f(y) - f((1-t)x + ty) \leq (1-t)f'((1-t)x + ty)(y-x).$$

Multiplicando a primeira desigualdade por $(1-t)$, a segunda por t e somando-as, obtemos:

$$f(x) - tf(x) - f((1-t)x + ty) + tf((1-t)x + ty) + tf(y) - tf((1-t)x + ty) \leq 0$$

$$(1-t)f(x) - f((1-t)x + ty) + tf(y) \leq 0$$

$$(1-t)f(x) + tf(y) \leq f((1-t)x + ty).$$

□

Podemos generalizar esses critérios para dimensões maiores.

A derivada primeira de $f(x) = f(x_1, x_2, ...x_n)$, para funções de várias variáveis, é uma transformação linear, cuja matriz na base canônica, chamada de matriz Jacobiana, é dada pelas derivadas parcias de primeira ordem de f:

$$Jf(x) = \left[\begin{array}{cccc} \frac{\partial f}{\partial x_1}(x) & \frac{\partial f}{\partial x_2}(x) & \cdots & \frac{\partial f}{\partial x_n}(x) \end{array} \right],$$

como já foi visto.

Teorema 14. Seja f uma função C^1 num subconjunto convexo A de $\mathbb{R}^n$. Então, f é côncava em A se, e somente se, para quaisquer x e $y \in A$.

$$f(y) - f(x) \leq f'(x).(y-x), \text{ ou seja,}$$

$$f(y) - f(x) \leq \left[\frac{\partial f}{\partial x_1}(x)(y_1 - x_1) + ... + \frac{\partial f}{\partial x_n}(x)(y_n - x_n)\right]$$

Da mesma forma, f é convexa em A se, e somente se,

$$f(y) - f(x) \geq f'(x).(y-x) \text{ para quaisquer } x \text{ e } y \in A.$$

Demonstração. Sejam x e y pontos arbitrários em A. Seja

$$g_{x,y} \equiv f(ty + (1-t)x)$$

$$g_{x,y} \equiv f(x + t(y-x))$$

$$g_{x,y} \equiv f(x_1 + t(y_1 - x_1), ..., x_n + t(y_n - x_n)).$$

Então, utilizando a regra da cadeia, temos:

$$g'_{x,y}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + t(y-x))(y_i - x_i).$$

Agora, considerando que f é côncava se, e somente se, cada uma destas $g_{x,y}$ é côncava, então como,

$$g_{x,y}(1) - g_{x,y}(0) \leq g'_{x,y}(0)(1-0)$$

$$g_{x,y}(1) - g_{x,y}(0) \leq g'_{x,y}(0).$$

Desta forma, para quaisquer x, $y \in A$,

$$f(y) - f(x) \leq f'(x)(y-x).$$

□

Corolário 2.1: Se f é uma função C^1 côncava em um conjunto convexo A e se $x_0 \in A$, então:

$$f'(x_0)(y-x_0) \leq 0 \Rightarrow f(y) \leq f(x_0).$$

Em particular, se $f'(x_0)(y-x_0) \leq 0$ para todo $y \in A$, então x_0 é um máximo global de f.

O teorema anterior é uma técnica muito útil para provar propriedades sobre funções côncavas e convexas, porém, por envolver a verificação de uma desigualdade para quaisquer x e y no domínio, em geral, não é um teste muito prático para verificar se uma função é côncava ou convexa. Utilizaremos, então uma generalização do teste da derivada segunda: f é côncava em um intervalo I se, e somente se, $f''(x) \leq 0$ para qualquer $x \in I$. A generalização natural da derivada segunda, $f''(x)$, de f real a valores reais para funções de várias variáveis é a matriz hessiana de todas as derivadas parciais de segunda ordem:

$$\mathscr{H}f(x) = \begin{bmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f_{x_nx_1} & f_{x_nx_2} & \cdots & f_{x_nx_n} \end{bmatrix}$$

onde escrevemos $f_{x_jx_i}$ para $\dfrac{\partial^2 f}{\partial x_i x_j}$ e cada variável é calculada no ponto x.

A generalização natural de $f''(x) \leq 0$ é a afirmação que a matriz hessiana $\mathscr{H}f(x)$ é não positiva em cada x do domínio de f. O teorema a seguir resume o teste da derivada segunda para funções côncavas e convexas em $\mathbb{R}^n$.

Teorema 15. Seja f uma função C^2 em um conjunto aberto convexo A de $\mathbb{R}^n$. Então, f é uma função côncava em A se, e somente se, a matriz hessiana $\mathscr{H}f(x)$ é negativa semidefinida para cada x em A. A função f é uma função convexa em A se, e somente se, $\mathscr{H}f(x)$ é positiva semidefinida para cada x em A.

Demonstração. Sejam x e y pontos quaisquer em A e $g_{x,y}(t) \equiv f(ty+(1-t)x)$. Então, f é côncava em A se, e somente se, cada $g_{x,y}(t)$ é côncava, que é equivalente a cada $g''_{x,y}(\mathrm{t}) \leq 0$. Assim, pelo teorema anterior e pela regra da cadeia, vimos que:

$$g'_{(x,y)}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+t(y-x))(y_i-x_i)$$

$$g'_{(x,y)}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x)(y_i-x_i)$$

$$g'_{(x,y)}(t) = f'(x)(y-x)$$

Logo,

$$g''_{x,y}(t) = \frac{\partial}{\partial t}\left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x+t(y-x))(y_i-x_i)\right)$$

$$g''_{x,y}(t) = \sum_{j=1}^{n}\sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x+t(y-x))(y_j-x_j)(y_i-x_i)$$

$$g''_{x,y}(t) = \sum_{i,j=1}^{n} (y_j-x_j)\frac{\partial^2 f}{\partial x_i \partial x_j}(x+t(y-x))(y_i-x_i)$$

$$g''_{x,y}(t) = (y-x)^T \mathscr{H} f(x+t(y-x))(y-x).$$

Se $\mathscr{H} f(z)$ é não positiva, então,

I) Cada $g''_{x,y}(t) \leq 0$;

II) Cada $g_{x,y}$ é côncava;

III) A própria f é côncava.

Reciprocamente, suponha que f é côncava em A. Seja z um ponto qualquer em A e seja v um vetor arbitrário em $\mathbb{R}^n$. Queremos mostrar que $v^T \mathscr{H} f(z)v \leq 0$. Como A é aberto, existe $t_0 > 0$ tal que $y = z + t_0 v$ está em A. Como f é côncava, $g_{x,y}$ é côncava e $g''_{z,y}(0) \leq 0$. Então,

$$0 \geq g''_{z,y}(0) = (y-z)^T \mathscr{H} f(z)(y-z)$$

$$= (t_0 v)^T \mathscr{H} f(z)(t_0 v)$$

$$=(t_0)^2[v^T \mathscr{H} f(z)v].$$

Assim, $v^T \mathscr{H} f(z)v \leq 0$ e $\mathscr{H} f(z)$ é não positiva para cada z em A.

□

Uma propriedade valiosa para o nosso estudo é que os pontos críticos das funções côncavas são automaticamente os máximos globais.

Teorema 16. Seja f uma função côncava (convexa) em um subconjunto aberto e convexo A de $\mathbb{R}^n$. Se x_0 é um ponto crítico de f, ou seja, se $\dfrac{\partial f}{\partial x_1}(x_0) = ... = \dfrac{\partial f}{\partial x_n}(x_0) = 0$, então $x_0 \in A$ é um máximo (mínimo) global de f em A.

Demonstração. Se f é côncava e $\frac{\partial f}{\partial x_1}(x_0) = ... = \frac{\partial f}{\partial x_n}(x_0) = 0$, então considerando que $f(y)$ - $f(x) \leq \frac{\partial f}{\partial x_1}(x)(y_1 - x_1)$ + ... + $\frac{\partial f}{\partial x_n}(x)(y_n - x_n)$, então $f(y) - f(x_0) \leq 0$ para $\forall$ $y \in A$. Ou seja, para $\forall y \in A$, temos que $f(y) \leq f(x_0)$, que significa que x_0 é um máximo global de f em A. $\square$

Dessa forma concluímos que:

I) Se x_0 é ponto crítico e $\mathscr{H}f(x)$ é não positiva em A, então f é côncava e x_0 é máximo global;

II) Se x_0 é ponto crítico e $\mathscr{H}f(x)$ é não negativa em A, então f é convexa e x_0 é mínimo global.

Pelo que vimos, o ”teste da derivada segunda” é mais fácil de ser aplicado, basta analisarmos se a matriz é positiva, negativa, não positiva, não negativa ou indefinida.

Temos a seguinte generalização do teorema anterior:

Teorema 17. Seja f uma função C^1 definida em um subconjunto convexo A de $\mathbb{R}^n$. Se f é uma função côncava e se x_0 é um ponto de A que satisfaz $f'(x_0)(y - x_0) \leq 0$ para $\forall$ $y \in A$, então $x_0 \in A$ é um máximo global de f em A. Se f é uma função convexa e se x_0 é um ponto de A que satisfaz $f'(x_0)(y - x_0) \geq 0$ para $\forall$ $y \in A$, então $x_0 \in A$ é um mínimo global de f em A.

Exemplo 32. Mostraremos que $f(x_1, x_2) = x_1^2 + x_2^2$ é convexa em $\mathbb{R}^2$.

Como vimos, uma função f é convexa se, e somente se, $f(y_1, y_2)$ - $f(x_1, x_2) \geq \frac{\partial f}{\partial x_1}(x_1, x_2)(y_1 - x_1) + \frac{\partial f}{\partial x_2}(x_1, x_2)(y_2 - x_2)$.

Assim temos,

$$(y_1 - x_1)^2 + (y_2 - x_2)^2 \geq 0$$

$$(y_1^2 - 2x_1y_1 + x_1^2) + (y_2^2 - 2x_2y_2 + x_2^2) \geq 0$$

$$y_1^2 + y_2^2 - x_1^2 - x_2^2 - 2x_1y_1 + 2x_1^2 - 2x_2y_2 + 2x_2^2 \geq 0$$

$$(y_1^2 + y_2^2) - (x_1^2 + x_2^2) \geq 2x_1y_1 - 2x_1^2 + 2x_2y_2 - 2x_2^2$$

$$(y_1^2 + y_2^2) - (x_1^2 + x_2^2) \geq 2x_1(y_1 - x_1) + 2x_2(y_2 - x_2).$$

O que é verdade para quaisquer (x_1, x_2) e $(y_1, y_2) \in \mathbb{R}^2$.

Outra forma de demonstrar esse resultado é observando que a matriz hessiana associada à f é dada por $\mathscr{H}f(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$, que é positiva para todo $(x,y) \in \mathbb{R}^2$.

Exemplo 33. Considerando a função $f(x,y) = x^4 + x^2y^2 + y^4 - 3x - 8y$ temos que sua matriz hessiana é dada por :

$$\mathscr{H}f(x,y) = \begin{bmatrix} 12x^2 + 2y^2 & 4xy \\ 4xy & 2x^2 + 12y^2 \end{bmatrix}, \text{ para } (x,y) \neq (0,0).$$

Dessa forma, os dois menores principais líderes, $12x^2 + 2y^2$ e $24x^4 + 148x^2y^2 + 24y^4$, são positivos, sendo assim, f é uma função convexa em todo $\mathbb{R}^2$.

Exemplo 34. Considerando a função Cobb-Douglas geral $U(x,y) = x^a y^b$, temos que sua matriz hessiana é dada por:

$$\mathscr{H}f(x,y) = \begin{bmatrix} a(a-1)x^{a-2}y^b & abx^{a-1}y^{b-1} \\ abx^{a-1}y^{b-1} & b(b-1)x^a y^{b-2} \end{bmatrix},$$

cujo determinante é

$$det\mathscr{H}f(x,y) = ab(1-a-b)x^{2a-2}y^{2b-2}.$$

Para que U seja côncava em $\mathbb{R}^2_+$, precisamos que $a(a-1) < 0$ e $ab(1-a-b) > 0$, ou seja, precisamos que $0 < a < 1$, $0 < b < 1$ e $a + b \leq 1$.

2.2 Otimização condicionada

2.2.1 Máximos e mínimos sobre conjunto compacto

Definição 13. Seja A um subconjunto de $\mathbb{R}^2$, dizemos que A é um conjunto limitado se A estiver contido em alguma bola aberta de centro na origem. Dizemos, por outro lado, que A é um conjunto fechado se o seu complementar $\left\{(x,y) \in \mathbb{R}^2/(x,y) \notin A\right\}$ for um conjunto aberto. Dizemos que A é um conjunto compacto se A for fechado e limitado.

Teorema 18. (Teorema de Weierstrass) Se f for contínua no compacto A, então existirão (x_1,y_1) e(x_2,y_2) em A tais que, para todo (x,y) em A $f(x_1,y_1) \leq f(x,y) \leq f(x_2,y_2)$, ou seja $f(x_1,y_1)$ é o valor mínimo e $f(x_2,y_2)$ é o valor máximo de f em A. [3]

Procedimento para obtenção de máximo e mínimo

O primeiro passo é encontrarmos os pontos críticos, pois são os únicos pontos interiores de A com possibilidade de serem extremantes. Para que isso seja feito, devemos supor que f admita derivadas parciais nos pontos interiores de A.

Em seguida, devemos determinar o máximo e o mínimo de f na fronteira de A.

Sendo assim, para obtermos o ponto máximo de f em A, comparamos os valores que f assume nos pontos críticos com o máximo valor de f na fronteira de A, o maior deles será o máximo de f em A. Da mesma forma, comparamos os valores que f assume nos pontos críticos com o mínimo valor de f na fronteira de A, o menor deles será o mínimo de f em A.

Exemplo 35. Uma determinada empresa está interessada em maximizar o lucro mensal proveniente de dois de seus produtos, designados I e II. Para fabricar esses produtos ela utiliza um tipo de máquina que tem uma disponibilidade de 200 máquinas-horas por mês e um tipo de mão de obra com uma disponibilidade de 240 homens-horas por mês. Para se produzir uma unidade do produto I utilizam-se 5 horas de máquina e 10 horas de mão de obra, enquanto para o produto II utilizam-se 4 horas de máquina e 4 horas de mão de obra. Espera-se uma demanda de 20 unidades por mês do produto I e 45 do produto II. Calcula-se um lucro, por unidade, de $R\$10,00$ para o produto I e $R\$6,00$ para o produto II. Determinaremos as quantidades de cada produto que deverão ser fabricadas por mês, para o lucro mensal ser máximo.

O problema consiste em maximizar o lucro $L = 10x + 6y$, onde x é a quantidade do produto I e y do produto II, com as restrições $0 \leq x \leq 20$, $0 \leq y \leq 45$, $5x + 4y \leq 200$, $10x + 4y \leq 240$.

Considerando o conjunto A, limitado pelos eixos coordenados x e y, e pelas retas: $L_1 = \{(x,y) \in \mathbb{R}^2 \ 0 \leq x \leq 6 \text{ e } y = 45\}$, $L_2 = \{(x,y) \in \mathbb{R}^2 \ 6 \leq x \leq 8 \text{ e } y = \dfrac{200 - 5x}{4}\}$, $L_3 = \{(x,y) \in \mathbb{R}^2 \ 8 \leq x \leq 20 \text{ e } y = \dfrac{240 - 10x}{4}\}$ e $L_4 = \{(x,y) \in \mathbb{R}^2 \ x = 20 \text{ e } 0 \leq y \leq 10\}$, vemos que L não possui pontos críticos nessa região e, analisando as extremidades de cada segmento de reta, encontramos $x = 8$ e $y = 40$ como máximo.

Exemplo 36. Seja $f(x,y) = xy$ restrita ao conjunto $A = \{(x,y) \in \mathbb{R}^2 \ x^2 + y^2 \leq 1\}$, temos que f assume valor máximo absoluto nos pontos $\left(\dfrac{\sqrt{2}}{2}, \dfrac{\sqrt{2}}{2}\right)$ e $\left(-\dfrac{\sqrt{2}}{2}, -\dfrac{\sqrt{2}}{2}\right)$, sendo o valor de f nesses pontos igual a $\dfrac{1}{2}$ e valor mímino absoluto nos pontos $\left(-\dfrac{\sqrt{2}}{2}, \dfrac{\sqrt{2}}{2}\right)$ e $\left(\dfrac{\sqrt{2}}{2}, \dfrac{\sqrt{2}}{2}\right)$, sendo o valor de f nesses pontos igual a $-\dfrac{1}{2}$.

De fato, analisando os pontos críticos, encontramos o ponto $(0,0)$, sendo que $f(0,0) = 0$, e portanto não é ponto de máximo nem de mínimo do problema.

Fazendo o estudo de f restrita à curva $x^2+y^2=1$, podemos escrever $x=\cos t$ e $y=\operatorname{sen} t$, $0\leq t\leq 2\pi$. Logo, f restrita à curva, é dada por:

$$g(t)=f(\cos t,\operatorname{sen} t)=\cos t.\operatorname{sen} t=\frac{\operatorname{sen}(2t)}{2},\ 0\leq t\leq 2\pi.$$

Para g, temos os seguintes pontos críticos: para $t=\frac{\pi}{4},\frac{3\pi}{4},\frac{5\pi}{4}$ e $\frac{7\pi}{4}$, que correspondem aos valores $\left(\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2},\frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$ e $\left(\frac{\sqrt{2}}{2},-\frac{\sqrt{2}}{2}\right)$, respectivamente.

Além disso, devemos analisar g nos extremos, isto é, para $t=0$ e $t=2\pi$, que correspondem ao ponto $(1,0)$. Logo, comparando f em todos esses valores, chegamos ao resultado.

2.2.2 Determinação de Candidatos a Extremantes Locais Condicionados pelo Método dos Multiplicadores de Lagrange

Seja f diferenciável no aberto A e seja $B=\{(x,y)\in A/g(x,y)=0\}$, onde g é suposta de classe C^1 em A. Suporemos, também, que $\nabla g(x,y)\neq(0,0)$ em B. Nosso objetivo é determinar uma condição necessária para que $(x_0,y_0)\in B$ seja um extremante local de f em B. Sem perda de generalidade, vamos supor que $(x_0,y_0)\in B$ seja um ponto de máximo local de f em B. Desta forma, existe uma bola aberta com centro em (x_0,y_0) que chamaremos de V, isto é,

$$f(x,y)\leq f(x_0,y_0),\ \forall (x,y)\in B\cap V.$$

Por meio do Teorema de Função Implícita, com os mesmos argumentos utilizados nas considerações geométricas do vetor gradiente, temos que existe uma curva $\vec{\gamma}$ diferenciável em um intervalo aberto I tal que $\vec{\gamma}(t_0)=(x_0,y_0)$, $t_0\in I$, $\vec{\gamma}'(t_0)\neq\vec{0}$ e $g(\vec{\gamma}(t))=0$, $\forall t\in I$.

Da continuidade da $\vec{\gamma}$ segue que $\exists\sigma>0$ / se $t\in(t_0-\sigma,t_0+\sigma)$, então $\vec{\gamma}(t)\in B\cap V$ (pequenas variações em t acarretam em pequenas variações de $\vec{\gamma}(t)$).

Logo, $f(\vec{\gamma}(t))\leq f(\vec{\gamma}(t_0))$, $\forall t\in(t_0-\sigma,t_0+\sigma)$ o que implica que t_0 é ponto de máximo local para $h(t)=f(\vec{\gamma}(t))$. Assim, $h'(t_0)=0$. Pela regra da cadeia,

$$h'(t)=\nabla f(\vec{\gamma}(t)).\vec{\gamma}'(t)$$

e segue, então, que

$$\nabla f(\vec{\gamma}(t_0)).\vec{\gamma}'(t_0)=0.$$

Por outro lado, como $g(\vec{\gamma}(t)) = 0$, novamente pela regra da cadeia,

$$\nabla g(\vec{\gamma}(t)).\vec{\gamma}'(t) = 0,$$

em particular,

$$\nabla g(\vec{\gamma}(t_0)).\vec{\gamma}'(t_0) = 0.$$

Como $\nabla g(\vec{\gamma}(t_0)) \neq \vec{0}$, segue que $\nabla f(x_0,y_0)$ e $\nabla g(x_0,y_0)$ são paralelos, isto é, existe $\lambda_0 \in \mathbb{R}/$

$$\nabla f(x_0,y_0) = \lambda_0 \nabla g(x_0,y_0).$$

Logo, se (x_0,y_0) é ponto de máximo local de f em B, (x_0,y_0) deve ser solução do sistema

$$\begin{cases} \nabla f(x,y) = \lambda \nabla g(x,y) \\ g(x,y) = 0 \end{cases},$$

para algum λ real.

Temos então, o seguinte teorema.

Teorema 19. Sejam f diferenciável no aberto $A \subset \mathbb{R}^2$ e $B = \{(x,y) \in A \; g(x,y) = 0\}$, onde g é suposta de classe C^1 em A, e $\nabla g(x,y) \neq (0,0)$, para todo $(x,y) \in B$. Uma condição necessária para que $(x_0,y_0) \in B$ seja extremante local de f em B é que exista um real λ_0 tal que:

$$\nabla f(x_0,y_0) = \lambda_0 \nabla g(x_0,y_0).$$

A variável λ_0 é denominada Multiplicador de Lagrange.

Se acrescentarmos as hipóteses $f \in C^1$ e $\nabla f(x_0,y_0) \neq 0$, podemos afirmar que a curva de nível de f tangencia a restrição $g(x,y) = 0$ no ponto (x_0,y_0).

Exemplo 37. Voltando ao exemplo anterior, podemos observar que para encontrar o valor máximo ou mínimo de f, restrita à circunferência de raio 1 e centro na origem, podemos utilizar este teorema usando $g(x,y) = x^2 + y^2 - 1 = 0$ como restrição.

Então, $\begin{cases} f(x,y) = xy \\ g(x,y) = x^2 + y^2 - 1 = 0 \end{cases}.$

Assim, $\begin{cases} \nabla f(x,y) = \lambda \nabla g(x,y) \\ g(x,y) = x^2 + y^2 - 1 = 0 \end{cases} \Rightarrow \begin{cases} (y,x) = \lambda(2x, 2y) \\ x^2 + y^2 - 1 = 0 \end{cases}$,

dessa forma $\lambda = \pm\dfrac{1}{2} \Rightarrow y = 0$ e $x = \pm 1$.

Verificamos que os pontos $(1,0)$, $(-1,0)$, $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ e $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ são candidatos a extremantes de f, restrita à circunferência. Como a própria circunferência é um conjunto compacto, f assume máximos e mínimos absolutos na curva, e são dados por $\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$, $\left(-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$ e $\left(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$.

A mesma análise pode ser feita para uma função de três variáveis.

Se considerarmos mais de uma restrição, temos o seguinte teorema:

Teorema 20. Sejam f diferenciável em $A \subset \mathbb{R}^3$ e $B = \{(x,y,z) \in A / g(x,y,z) = 0$ e $h(x,y,z) = 0\}$, onde g e h são supostamente de classe C^1 em A e $\nabla g(x,y,z) \wedge \nabla h(x,y,z) \neq \vec{0}$ em B. Nessas condições, para que $(x_0,y_0,z_0) \in B$ seja extremante local de f em B, existem λ_1 e $\lambda_2 \in \mathbb{R}$ tal que $\nabla f(x_0,y_0,z_0) = \lambda_1 \nabla g(x_0,y_0,z_0) + \lambda_2 \nabla h(x_0,y_0,z_0)$.

Demonstração. Suponhamos que (x_0,y_0,z_0) seja ponto de máximo local de f em B tal que exista uma bola aberta D de centro $(x_0,y_0,z_0) / \forall (x,y,z) \in B \cap D$, $f(x,y,z) \leq f(x_0,y_0,z_0)$.

Consideremos uma curva diferenciável $\gamma : I \to \mathbb{R}^3$, I aberto, tal que $\gamma(t_0) = (x_0,y_0,z_0)$, $\gamma'(t_0) \neq \vec{0}$ e $\gamma(t) \in B$ para todo t em I. Como γ é contínua, existe $\phi > 0$ tal que $t \in]t_0 - \phi, t_0 + \phi[\Rightarrow \gamma'(t) \in B \cap D$.

Assim, para todo $t \in]t_0 - \phi, t_0 + \phi[$ temos $f(\gamma(t)) \leq f(\gamma(t_0))$.

Logo, t_0 é ponto de máximo local de $F(t) = f(\gamma(t))$ e daí $F'(t_0) = 0$, ou seja, $\nabla f(\gamma(t_0)).\gamma'(t_0) = 0$.

Por outro lado, se $\gamma(t) \in B$ para todo $t \in I$, temos $g(\gamma(t)) = 0$ e $h(\gamma(t)) = 0$, assim $\nabla g(\gamma(t_0)).\gamma'(t_0) = 0$ e $\nabla h(\gamma(t_0)).\gamma'(t_0) = 0$.

Como o produto escalar de $\gamma'(t_0)$ com cada um dos gradientes da função dada e das funções restrições é zero, temos que o gradiente da função dada está contido no plano determinado pelos gradientes das restrições, dessa forma o gradiente de f é uma combinação linear dos gradientes de h e g, ou seja, existem reais λ_1 e λ_2, tais que

$$\nabla f(\gamma(t_0)) = \lambda_1 \nabla g(\gamma(t_0)) + \lambda_2 \nabla h(\gamma(t_0)).$$

Então, os candidatos a extremantes locais de f em B são os pontos $(x,y,z) \in A$ que tornam possível o sistema:

$$\begin{cases} \nabla f(x,y,z) & = & \lambda_1 \nabla g(x,y,z) + \lambda_2 \nabla h(x,y,z) \\ g(x,y,z) & = & 0 \\ h(x,y,z) & = & 0 \end{cases}$$ □

Vejamos um exemplo sobre a teoria.

Exemplo 38. Um ginásio poliesportivo será construído de acordo com as seguintes condições:

I) Deve existir um jardim com formato elíptico interceptado por duas pistas perpendiculares no centro do jardim.

II) As pistas dividirão a parte interna ao jardim em 4 partes de igual área, onde serão localizadas as quadras.

III) Será construído um muro tangente ao jardim e o portão de acesso ao ginásio se localizará no ponto de tangência.

IV) O objetivo é construir o muro de forma que a região delimitada por ele e pelas pistas perpendiculares tenha a menor área possível e que se encontre no primeiro quadrante em relação às pistas.

Considerando que a equação da elipse que representará o jardim é dada por $4x^2 + y^2 = 4$, determine a localização do portão de entrada em relação ao centro do jardim e a equação da reta que representa o muro a ser construído.

Resolução:

Temos a seguinte situação: devemos determinar a equação da reta tangente a elipse $4x^2 + y^2 = 4$, ou seja, $y = 2\sqrt{1 - x_0^2}$ no ponto $(x_0, y_0) = (x_0, f(x_0))$, dada por:

$$\begin{aligned} y &= f'(x_0)(x - x_0) + f(x_0) \\ y &= \frac{-2x_0}{\sqrt{1 - x^2}}(x - x_0) + 2\sqrt{1 - (x_0)^2} \\ y &= \frac{-2x_0 x + 2(x_0)^2 + 2 - 2(x_0)^2}{\sqrt{1 - (x_0)^2}} \end{aligned}$$

$$y = \frac{2(-x_0 x + 1)}{\sqrt{1 - (x_0)^2}}.$$

Como precisamos do triângulo determinado por essa reta tangente e pelos eixos coordenados (pistas perpendiculares), precisamos determinar os pontos de intersecção dessa reta com esses eixos.

I) para $y = 0$, temos:

$$\frac{2(-x_0a+1)}{\sqrt{1-(x_0)^2}} = 0$$
$$\Rightarrow 2(-x_0a+1) = 0$$
$$\Rightarrow ax_0 = 1$$
$$\Rightarrow a = \frac{1}{x_0},$$

ou seja, obtemos o ponto $\left(\frac{1}{x_0}, 0\right)$.

II) para $x = 0$, temos:

$$\frac{2.2(-x_0.0+1)}{2\sqrt{1-(x_0)^2}} = b$$
$$\Rightarrow \frac{4}{y_0} = b$$
$$\Rightarrow b = \frac{4}{y_0},$$

e portanto, o ponto de intersecção é dado por $\left(0, \frac{4}{y_0}\right)$.

Dessa forma a área do triângulo é dada por:

$$A = \frac{ab}{2} = \frac{\left(\frac{1}{x_0}\right)\left(\frac{4}{y_0}\right)}{2} = \frac{2}{x_0y_0}.$$

Além disso está sujeita à restrição

$$g(x_0, y_0) = 4x_0^2 + y_0^2 - 4.$$

Como a área deve ser mínima, devemos ter:

$$\begin{cases} \nabla A(x_0, y_0) = \lambda g(x_0, y_0) \\ g(x_0, y_0) = 0. \end{cases}$$

O sistema acima é equivalente a

$$\begin{cases} \frac{\partial A}{\partial x_0}(x_0, y_0) = \lambda \frac{\partial g}{\partial x_0}(x_0, y_0) \\ \frac{\partial A}{\partial y_0}(x_0, y_0) = \lambda \frac{\partial g}{\partial y_0}(x_0, y_0) \\ 4x_0^2 + y_0^2 - 4 = 0 \end{cases}$$

o qual implica em

$$\begin{cases} -\dfrac{2}{x_0^2 y_0} = 8\lambda x_0 \\ -\dfrac{2}{y_0^2 x_0} = 2\lambda y_0 \\ 4x_0^2 + y_0^2 = 4. \end{cases}$$

A primeira equação acima nos fornece:

$$\lambda = \frac{-1}{4x_0^3 y_0}.$$

Quanto à segunda, temos

$$\frac{-2}{y_0^2 x_0} = \left(\frac{-1}{x_0^3 y_0}\right)\frac{y_0}{2} \Rightarrow y_0^2 x_0 = 4x_0^3 \Rightarrow y_0^2 x_0 - 4x_0^3 = 0 \Rightarrow x_0(y_0^2 - 4x_0^2) = 0,$$

donde segue que $x_0 = 0$ ou

$$4x_0^2 = y_0^2 \Rightarrow x_0^2 = \frac{y_0^2}{4} \Rightarrow x_0 = \pm\frac{y_0}{2}.$$

Para $x_0 = 0$,

$$\frac{y_0^2}{4} = 1 \Rightarrow y_0^2 = 4 \Rightarrow y_0 = \pm 2$$

e para $x_0 = \pm\dfrac{y_0}{2}$,

$$\frac{y_0^2}{4} + \frac{y_0^2}{4} = 1 \Rightarrow 2y_0^2 = 4 \Rightarrow y_0^2 = 2 \Rightarrow y_0 = \pm\sqrt{2}.$$

Como $x > 0$ e $y > 0$, o ponto que nos interessa é $\left(\dfrac{\sqrt{2}}{2}, \sqrt{2}\right)$ e a equação da reta tangente à elipse dada é $y = 2\dfrac{(-x_0 x + 1)}{\sqrt{1 - x_0^2}}$

Com $x_0 = \dfrac{\sqrt{2}}{2}$,

$$y = \frac{2\left(-\dfrac{\sqrt{2}}{2}x + 1\right)}{\sqrt{1 - \left(\dfrac{\sqrt{2}}{2}\right)^2}}$$

$$\Rightarrow y = \frac{-\sqrt{2}x + 2}{\sqrt{1 - \dfrac{1}{2}}}$$

$$\Rightarrow y = \frac{-\sqrt{2}x + 2}{\frac{1}{\sqrt{2}}}$$

$$\Rightarrow y = (-\sqrt{2}x + 2)\sqrt{2}$$

$$\Rightarrow y = -2x + 2\sqrt{2}.$$

CAPÍTULO

3

CONSIDERAÇÕES SOBRE PROBLEMAS DE OTIMIZAÇÃO NO ENSINO BÁSICO

A teoria desenvolvida nos capítulos anteriores é, normalmente, abordada em uma segunda disciplina de Cálculo Diferencial, em diversos cursos de exatas. Em alguns cursinhos ou para turmas mais ”interessadas em Matemática” é possível definir o conceito de derivada e aplicações em problemas de otimização que são bastante motivadores. Neste contexto, os problemas apresentados durante o trabalho poderiam ser desenvolvidos com essas turmas.

Entretanto, acreditamos ser possível trabalhar os conceitos de máximos e mínimos de uma função, sem citar explicitamente o conceito de derivação, instigando os alunos com problemas práticos e motivadores, questionando como abordá-los em situações onde o conceito de derivação faz-se necessário.

Listamos abaixo alguns exemplos onde podemos trabalhar este tema e formas de abordá-los.

1) Problemas em que a função a ser otimizada é uma função do segundo grau. Dois problemas clássicos neste sentido são descritos a seguir.

I) Um fazendeiro deve cercar dois pastos retangulares de dimensões a e b, com um lado comum a. Se o fazendeiro possui 9600 m de cerca, determinar as dimensões de a e b para que cada pasto tenha a máxima área possível.

Como os dois terrenos possuem um lado a comum, a quantidade de cerca utilizada é dada por $3a+4b$ e a área de cada pasto é representada por $a.b$.

Dessa forma temos:

$$\begin{cases} 3a+4b=9600 \\ A=a.b \end{cases}$$

Considerando b em função de a podemos escrever $b = 2400 - \frac{3}{4}a$, assim $A = 2400a - \frac{3}{4}a^2$ e a área máxima ocorre no vértice da parábola que representa o gráfico de uma função do segundo grau, nesse caso A é dada em função de a.

Assim, o valor de a para que a área seja máxima é o x do vértice da parábola $A = 2400a - \frac{3}{4}a^2$.

As raízes da equação $2400a - \frac{3}{4}a^2 = 0$ são $a = 0$ e $a = 3200$ e o valor de x do vértice é igual à média das raízes, como podemos verifcar no Figura 17 a seguir.

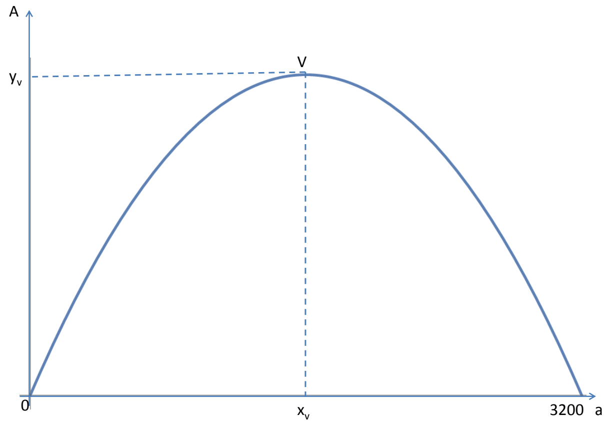

Figura 17 – Área máxima

Então $x_{vertice} = 1600$ metros. Dessa forma, podemos concluir que as dimensões para que a área seja máxima são $a = 1600$ metros e $b = 1200$ metros.

II) (retirado de [5],p.184) Um galpão retangular deve ser construído em um terreno com a forma de um triângulo retângulo com catetos medindo 20 metros e 10 metros, sendo que dois dos lados do galpão devem estar contidos nesses dois lados do terreno e o quarto vértice do retângulo que representa o galpão deve pertencer à hipotenusa do terreno, como está representado na Figura 18.

Vamos deteminar a área máxima desse galpão.

Considerando um plano cartesiano, onde os eixos coordenados são as retas que contém os catetos do terreno e x e y as dimensões do galpão, conforme Figura 19.

Observando a figura, temos que a área do galpão é dada por $A = xy$ e que o ponto P deve estar sobre a reta $x + 2y = 20$, então podemos expressá-lo como $P\left(x, 10 - \frac{x}{2}\right)$.

Dessa forma, podemos escrever a área como função de x da seguinte maneira:

$$A(x) = 10x - \frac{x^2}{2}.$$

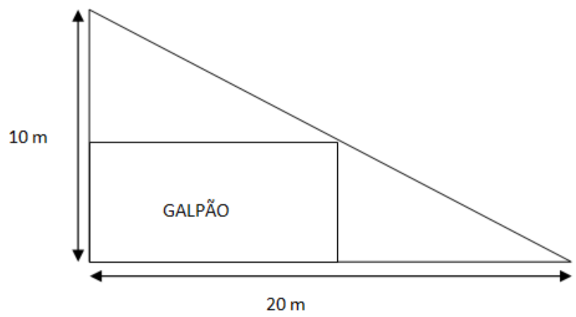

Figura 18 – Dimensões do terreno onde será construído o galpão

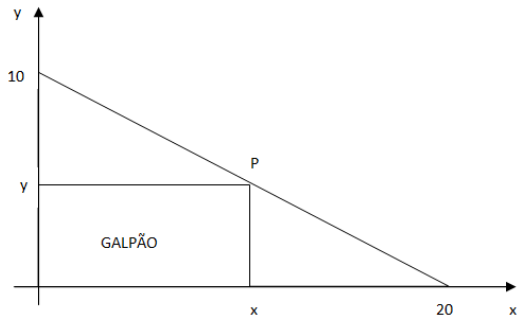

Figura 19 – Plano cartesiano e variáveis utilizados no cálculo do terreno de área máxima

Como vimos no primeiro problema, a área máxima ocorre no vértice da parábola que representa o gráfico de uma função do segundo grau.

Assim, o valor de x para que a área seja máxima é o x do vértice da parábola $A(x) = 10x - \frac{x^2}{2}$.

As raízes da equação $10x - \frac{x^2}{2} = 0$ são $x = 0$ e $x = 20$ e o valor de x do vértice é igual à média das raízes, como podemos verifcar na Figura 20.

Então,

$$x_{vertice} = 10 \text{ metros.}$$

Nosso objetivo é determinar a área máxima para o galpão e como isso ocorre no vértice da parábola, o valor que nos interessa é a imagem de x do vértice.

Sendo assim,

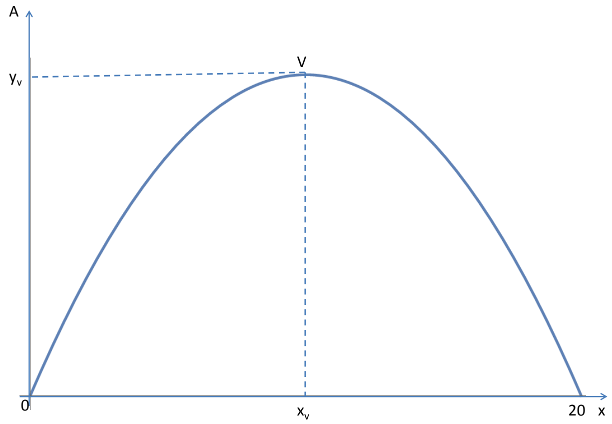

Figura 20 – Área Máxima

$$A_{maxima} = 50m^2 \ .$$

2)(retirado de [4], p.218 e 219) Problemas em que a utilização de softwares podem ajudar na interpretação (por exemplo, a função a ser otimizada é uma função polinomial do 3o. grau).

Em Biologia, encontramos a fórmula $\phi = VA$, onde ϕ é o fluxo de ar na traqueia, V é a velocidade do ar e A a área do círculo formado ao seccionarmos a traqueia.

Quando tossimos, o raio diminui, aumentando a velocidade do ar na traqueia. Sendo r_0 o raio normal da traqueia, a relação entre a veocidade V e o raio r da traqueia durante a tosse é dada por $V(r) = ar^2(r_0 - r)$, onde a é uma constante positiva. Podemos representar essa situação graficamente como na Figura 21.

Vamos calcular o raio r em que é maior a velocidade do ar e o valor de r com o qual teremos o maior fluxo possível.

O raio r da traqueia contraída, não pode ser maior que o raio normal r_0, nem menor que zero, ou seja, $0 \leq r \leq r_0$.

Precisamos determinar o máximo absoluto da função $V(r)$ em que $0 \leq r \leq r_0$.

Temos:

$$V(r) = ar^2(r_0 - r) \Rightarrow V'(r) = 2ar_0r - 3ar^2.$$

Fazendo $V'(r) = 0$, obtemos $r_1 = \dfrac{2}{3}r_0$ e $r_2 = 0$.

Temos $V''(r) = 2ar_0 - 6ar$. Como $V''(0) = 2ar_0 > 0 \Rightarrow r_2 = 0$, que dessa forma é um mínimo relativo. Como $V"(r) = \dfrac{2}{3}r_0$ é um valor negativo, concluímos que $r_1 = \dfrac{2}{3}r_0$ é um valor máximo relativo.

Para $r \in [0, r_0]$, temos que o máximo absoluto é $V\left(\frac{2}{3}r_0\right) = \frac{4a}{27}r_0^3$.

Diante desse resultado, afirmamos que a velocidade do ar na traqueia é maior quando o raio dela é dois terços do raio r_0 da traqueia não contraída.

Agora, determinaremos o máximo absoluto da função $\phi(r) = ar^2(r_0 - r)\pi r^2$.

Temos $\phi'(r) = 4a\pi r_0 r^3 - 5a\pi r^4$.

Fazendo $\phi'(r) = 0$, obtemos $r_1 = 0$ e $r_2 = \frac{4}{5}r_0$ como pontos críticos de $\phi(r)$.

Temos $\phi''(r) = 12a\pi r_0 r^2 - 20a\pi r^3$.

Logo, $\phi''(0) = 0$ e ϕ" $\left(\frac{4}{5}r_0\right) = -\frac{64}{25}a\pi r_0^3$ e concluímos que em $\frac{4}{5}r_0$ temos um ponto de máximo relativo.

O ponto $r_1 = 0$ é um ponto de mínimo relativo, pois a função $\phi(r)$ decresce em $(-\infty, 0]$ e cresce em $\left[0, \frac{4}{5}r_0\right]$.

O máximo absoluto em $[0, r_0]$ será $\phi\left(\frac{4}{5}r_0\right)$, que é igual a $\frac{256}{3125}a\pi r_0^5$.

Portanto, o maior fluxo possível é obtido quando $r = \frac{4}{5}r_0$.

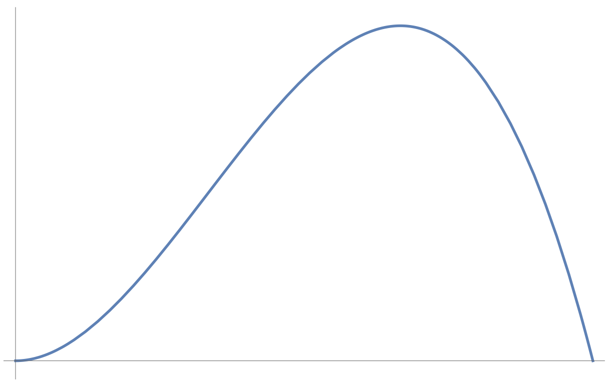

Figura 21 – Velocidade do ar na traquéia em função do raio da traquéia durante a tosse

3)(retirado de [6], p.33) Problemas envolvendo desigualdades, médias aritmética e geométrica

Considerando o paralelepípedo retângulo de dimensões x, y e z, representado na Figura 22, temos que sua área e seu volume são dados, respectivamente, por $A = 2xy + 2xz + 2yz$ e $V = xyz$.

Determinaremos os valores que minimizam a função $A(x, y, z) = 2xy + 2xz + 2yz$ dado que $V = xyz$ seja constante.

Assim, sabendo que a média aritmética entre n valores é maior que a média geométrica dos mesmos valores, temos que :

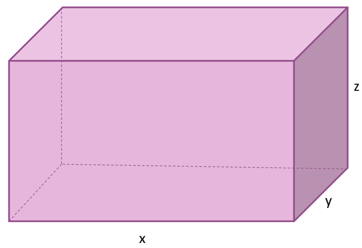

Figura 22 – Paralelepípedo

$$\frac{2xy+2xz+2yz}{3} \geq \sqrt[3]{2xy.2xz.2yz} \Rightarrow 2xy+2xz+2yz \geq 3.2\sqrt[3]{x^2y^2z^2} \Rightarrow A \geq 6\sqrt[3]{V^2}.$$

Como V é constante, a desigualdade entre as médias aritmética e geométrica nos diz que o menor valor que A pode assumir é $A = 6\sqrt[3]{V^2}$, assim :

$A = 2(xy+xz+yz) = 6\sqrt[3]{V^2}$.

Logo, $xy = xz = yz \Rightarrow x = y = z$, assim fazendo $z = x$ e $y = x$, temos $x = \sqrt[3]{V}$.

Podemos concluir, então que o paralelepípedo retângulo procurado é um cubo de aresta $\sqrt[3]{V}$.

4)(retirado de [6], p.34) Problemas analisando a região de otimização.

Um fabricante de ração animal dispõe de 60 sacos de milho e 32 sacos de trigo para produzir dois tipos de rações A e B. A ração A é composta de 3 sacos de milho, 4 sacos de trigo e gera um lucro de $R\$20,00$. A ração B é composta de 6 sacos de milho, 2 sacos de trigo e gera um lucro de $R\$24,00$. Determinaremos quanto de cada ração deve ser produzido para que o lucro do fabricante seja o maior possível.

Representaremos os dados do problema dispostos na Tabela 1 para facilitar a compreensão do mesmo.

Consideraremos x e y as quantidades de rações produzidas dos tipos A e B, respectivamente.

O lucro é obtido através da função $L(x,y) = 20x+24y$, diante das seguintes condições:

$$3x+6y \leq 60 \text{ , } 4x+2y \leq 32, \; x \geq 0 \text{ e } y \geq 0.$$

O nosso objetivo é estabelecer para quais valores de x e y a função lucro alcançará seu valor máximo, obviamente respeitadas as restrições acima.

Figura 23 – Tabela 1: Informações sobre a produção de ração animal

	Lucro (em R$)	Milho (em sacos)	Trigo (em sacos)
Ração (A)	20	3	4
Ração (B)	24	6	2
Disponível		60	32

Observando que todas as expressões relacionadas ao problema são lineares e que dependem de duas variáveis, x e y, podemos representar cada uma dessas expressões no plano cartesiano e analisar seu comportamento (ver Figuras 24 e 25).

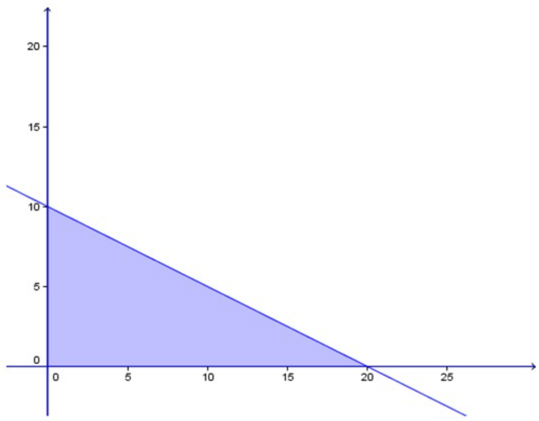

Figura 24 – Restrição 1, $3x+6y \leq 60$

A região hachurada representa os pontos que satisfazem cada uma das restrições. Logo, para que todas as restrições sejam satisfeitas simultaneamente fazemos a intersecção, como representado na Figura 26.

Chamaremos de região factível a região hachurada mais escura do plano cartesiano, e a mesma representa os pontos (x, y) que satisfazem as condições do problema.

Para que o valor de $L(x,y)$ seja máximo, devemos ter os maiores valores possíveis para x e y que satisfaçam as restrições. Traçando retas paralelas aos eixos coordenados, percebemos que tal ponto ocorre no vértice da região factível, como podemos acompanhar na Figura 27.

Portanto, para encontrar tal ponto devemos resolver o seguinte sistema:

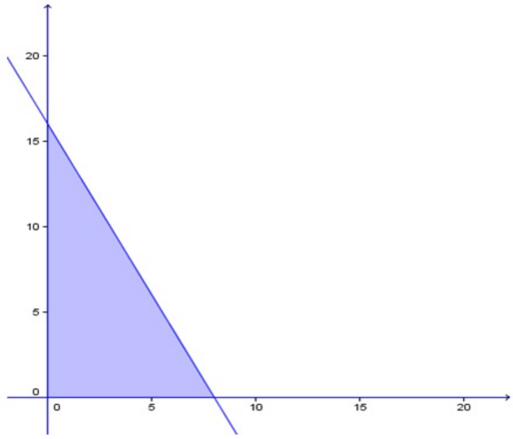

Figura 25 – Restrição 2, $4x+2y \leq 32$

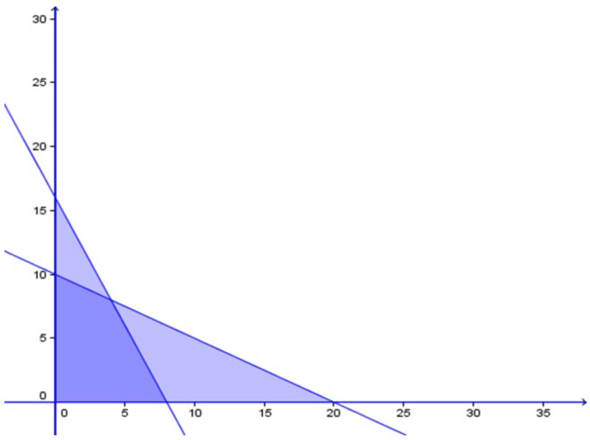

Figura 26 – Intersecção entre as restrições 1 e 2, $3x+6y \leq 60 \cap 4x+2y \leq 32$

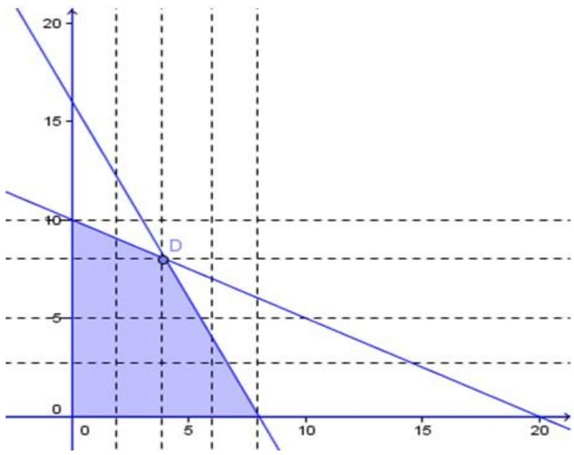

Figura 27 – Região factível e retas paralelas aos eixos coordenados

$$\begin{cases} 3x+6y=60 \\ 4x+2y=32 \end{cases},$$

cuja solução é o par ordenado $(4,8)$.

Assim, temos que o vértice do polígono é o ponto $(4,8)$, logo $x=4$ e $y=8$ são os valores que maximizam a função $L(x,y)=20x+24y$.

Portanto as quantidades das rações dos tipos A e B são, respectivamente, iguais a 4 e 8, e dessa forma o lucro máximo será igual a $R\$272,00$.

REFERÊNCIAS

[1] GUIDORIZZI, H.L. *Um curso de cálculo.* v. 2, 5. ed. Rio de Janeiro: LTC, 2014. Citado nas páginas 21 e 42.

[2] SIMON; CARL.D; LAWRENCE. *Matemática para economistas. Tradução: Claus Ivo Doring* Porto Alegre: Bookman, 2004. Citado nas páginas 57, 65, 68 e 73.

[3] LIMA, E.L. *Análise Real: Funções de várias variáveis.* v. 2, 12. ed. Rio de Janeiro-RJ: Coleção Matemática Universitária: IMPA, 2014. Citado nas páginas 41, 51 e 81.

[4] FLEMMING; GONÇALVES *Cálculo A: Funções, limite, derivação e integração.* 6 ed. São Paulo-SP: PEARSON, 2010. Citado na página 94.

[5] FLEMMING; GONÇALVES *Cálculo B: Funções de várias variáveis, integrais múltiplas, integrais curvilíneas e de superfície.* 2 ed. São Paulo-SP: PEARSON, 2009. Citado na página 92.

[6] COIMBRA *Resolução de Problemas do Ensino Médio Usando Métodos de Otimização.* Dissertação(Mestrado em Matemática) Universidade Federaldo Piauí, Teresina, 2013. Citado nas páginas 95 e 96.